Amit Kumar
Stefan Mozar
Jan Haase *Editors*

Advances in Cognitive Science and Communications

Selected Articles from the
5th International Conference on
Communications and Cyber-Physical
Engineering (ICCCE 2022), Hyderabad,
India

Springer

Cognitive Science and Technology

Series Editors

David M. W. Powers, College of Science and Engineering, Flinders University at Tonsley, Tonsley, SA, Australia

Richard Leibbrandt, College of Science and Engineering, Flinders University at Tonsley, Tonsley, SA, Australia

This series aims to publish work at the intersection of Computational Intelligence and Cognitive Science that is truly interdisciplinary and meets the standards and conventions of each of the component disciplines, whilst having the flexibility to explore new methodologies and paradigms. Artificial Intelligence was originally founded by Computer Scientists and Psychologists, and tends to have stagnated with a symbolic focus. Computational Intelligence broke away from AI to explore controversial metaphors ranging from neural models and fuzzy models, to evolutionary models and physical models, but tends to stay at the level of metaphor. Cognitive Science formed as the ability to model theories with Computers provided a unifying mechanism for the formalisation and testing of theories from linguistics, psychology and philosophy, but the disciplinary backgrounds of single discipline Cognitive Scientists tends to keep this mechanism at the level of a loose metaphor. User Centric Systems and Human Factors similarly should inform the development of physical or information systems, but too often remain in the focal domains of sociology and psychology, with the engineers and technologists lacking the human factors skills, and the social scientists lacking the technological skills. The key feature is that volumes must conform to the standards of both hard (Computing & Engineering) and social/health sciences (Linguistics, Psychology, Neurology, Philosophy, etc.). All volumes will be reviewed by experts with formal qualifications on both sides of this divide (and an understanding of and history of collaboration across the interdisciplinary nexus).

Indexed by SCOPUS

Amit Kumar · Stefan Mozar · Jan Haase
Editors

Advances in Cognitive Science and Communications

Selected Articles from the 5th International Conference on Communications and Cyber-Physical Engineering (ICCCE 2022), Hyderabad, India

Springer

Editors
Amit Kumar
Bioaxis DNA Research Centre Private Ltd.
Hyderabad, Telangana, India

Stefan Mozar
Glenwood, NSW, Australia

Jan Haase
Nordakademie University of Applied
Sciences
Elmshorn, Schleswig-Holstein, Germany

Contents

A Methodology for Selecting Smart Contract in Blockchain-Based Applications

Rohaila Naaz and Ashendra Kumar Saxena

Abstract In the era of e-commerce and digital economy where business parties as well as consumer-provider relationship are largely built on their agreement between them, agreement could comprise of several terms and policies on which both the parties are obligated, but in blockchain-based application, these terms and conditions are written in smart contract; smart contract is the executable code which can be executed in hardware or software without manual intervention. This paper reviews various practical applications of blockchain and challenges faced in implementing smart contract over blockchain. Second contribution of this paper is, we are proposing a methodology to be followed for designing and implementing smart contract in context of legal policies and difficulties encountered, as of right now there is no universal legal framework to be followed for smart contracts.

Keywords Blockchain applications · Smart contract · Methodology · Use case

1 Introduction

Smart contract is an executable code which is first introduced by Nick Szabo in 1994, this is an executable code built with the blockchain transactions to deploy terms and conditions for that same transaction over the blockchain network, the use of smart contract enables the application to run these transactions smoothly without any manual intervention required, and this removal of manual intervention further eliminates the need of third-party validators. Smart contracts also have the ability to run on required hardware and software, when blockchain is concerned. The very first application of blockchain, i.e., "Bitcoin" is introduced by Satoshi Nakamoto. It does not have smart contract feature but Ethereum platform which is a modified version of Nakamoto consensus which has the ability of deploying smart contracts. The basic structure of smart contract is shown in Fig. 1.

R. Naaz (✉) · A. K. Saxena
Faculty of Engineering and Computing Sciences, TMU, Moradabad, India
e-mail: rohila.computers@tmu.ac.in

1
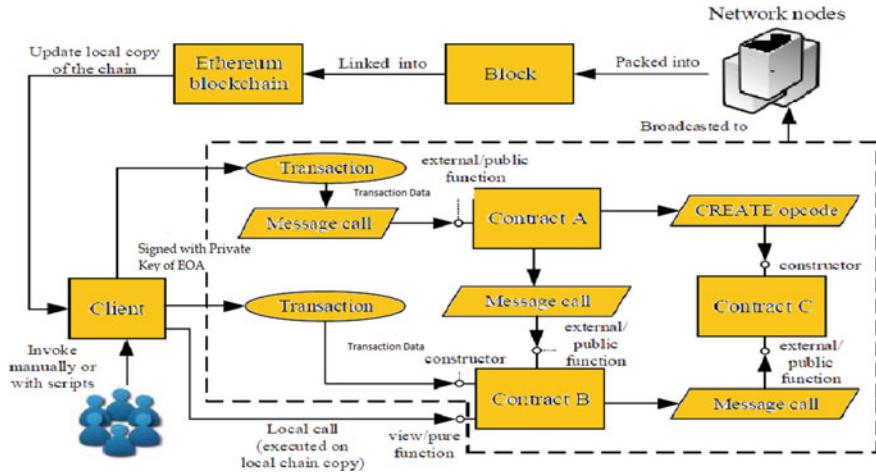
Fig. 1 Basic structure of smart contract [1]

Generally, languages for writing smart contracts are Solidity, Go, Python, Java, and Simplicity and ran over and platforms like Ethereum, Ripple, Quorum, Hyperledger Fabric, and R3 Corda.

Suppose if you want to sell an asset to someone, then you can simply use a smart contract on blockchain with your asset information, information about the asset could be stored in blockchain, and anyone can see this information who are on blockchain network, when one wants to buy your asset, then you can run a smart contract which could specify the terms like how much price the asset has been sold, if there is added transportation charges if required, handling ownership of that asset, etc., when this smart contract gets executed, then ownership has been transferred to the buyer, and this information is again visible to the users of the network, in that way blockchain maintains the transparency of the assets ownership. For a wide range of potential applications, blockchain-based smart contracts could offer a number of benefits such as speed and real-time updates, accuracy, lower execution risk, fewer intermediaries, lower cost, new business, or operational models.

2 Background

Various researches on use cases and their smart contracts have been published such as in [2], they have used different use cases like supply chain, Internet of Things, insurance, financial system, etc. and show how the smart contract can be in-built to network, and when it is integrated, the transaction will automatically execute the smart contract, but there is a research gap they have not mentioned that blockchain is either public or permissioned-based, that if the blockchain is permissioned-based

blockchain, then the participants who are involved in the transaction should be able to see and validate that transaction, and it is more secure and trustful.

Lopez vizar et al. analyze smart contracts from the legal perspective and contract law. Alberti et al. discusses JSON platform based on AgentSpeak language and the belief desire intention (BDI) model for self aware MAS implementation. Bordini et al. [3] survey programming languages for MAS implementation. Celaya et al. [4] discuss the abstract model of the MAS architecture and the evaluation of the coordination protocols in MAS with Petri nets.

Mahunnah et al. [5] discuss an agent-oriented model (AOM) and behavior-interface model for capturing socio-technical system behavior and capture the advantages of colored Petri nets (CPN) model over AOM for tool-supported simulation, model checking, and performance testing-based verification. The state of the art shows that the current smart contract approaches lack a required degree of intelligent automation to provide self-awareness and human-readability in smart contracts for their legal enforceability [2]. Rekin et al. [6] give the concept of weak smart contract and strong smart contract on the basis of the cost required to alter the smart contract means if the court or third party requires less cost to alter the smart contract then it is weak and if it requires an instance amount of cost then it is strong smart contract, and they also define smart contract in public permissionless blockchain in terms of consensus that if majority is saying it is true then it is but as it also gives the limitation if we want to deploy permission-based private blockchain where authenticity of masses not matters but verification from all the parties involved does matter.

Sabatucci et al. [7] discuss belief-desire-intention model [8] for making smart contracts self-aware, and in BDI model, racial planning on user belief, desire, and intention play a major role of making plans, so that an AI-based approach could be deployed on smart contracts for making them self-aware to make sure flexibility of the smart contracts or in order to make strong smart contracts [9].

Paper [10] discusses about the hybrid approach of the blockchain, it gives the idea about on-blockchain and off-blockchain methods with the use case of data seller and buyer of the network, it uses private Ethereum Turing complete Solidity language for writing smart contract and suggests master slave model where master node will work on off-blockchain and slave node will work on-blockchain, and it mainly addresses to reduce the complexity and computation required by the blockchain over vast network.

The main problem occurs with the validity of the smart contracts because there is a particularity of smart contracts deployed on-blockchain is that because of their decentralization and openness, and they are hard to amend after deployment. Therefore, we suggest that smart contracts are thoroughly validated (e.g., using conventional model checking tools) to uncover potential logical inconsistencies of their clauses (omissions, contradictions, duplications, etc.). In addition, we suggest that the actual implementation be systematically tested before deployment.

Paper [11] discusses about the languages of smart contract, they find the gaps that mostly smart contracts are written in Solidity language which is Turing complete, but Solidity is not flexible in terms of ontological expressiveness of business contracts and legal perspective of smart contract execution, so they are proposed to design a smart contract language whose verifiability would be given by using ANTLR Tool.

3   Smart Contract Use Cases

There are various use cases of smart contract, we will discuss its working, and to find the limitations and challenges faced, we execute smart contract on these specific applications.

A.   *Distributed Energy Transaction Mechanism*

Yu et al. [12] discuss the Chinese power distribution market. Nowadays, there are various forms of energy, be it electrical energy, solar energy, hydraulic energy. The smart contract will work on the consortium-based blockchain network of prosumers which executes transactions between peer-to-peer nodes. The distributed energy transaction flow is shown in Fig. 2, while the functions which are in-built on the smart contract are as shown in Fig. 3.

The main limitation on this p2p approach is how to localize the prosumers on the trading platform if the number of prosumers would get increased, as blockchains have network scalability problem.

B.   *Implementing a Microservices System*

Microservices system is a software engineering concept in which microservices allow teams to slice existing information systems into small and independent services that can be developed independently by different teams [13].

C.   *Secure Electric Vehicles Charging in Smart Community*

Yuntao et al. [14] propose a framework for the electric chargeable vehicles smart community, and they done it by introducing an energy blockchain which is a permissioned-type blockchain integrated with RES, means that only verified users can be a part of the network and transactions are to be validated by only the participating parties. After that based on the contract theory, the optimal contracts are

Fig. 2 Distributed energy transaction flow
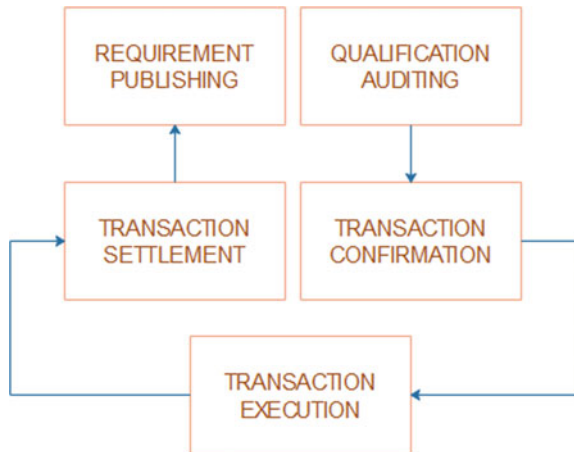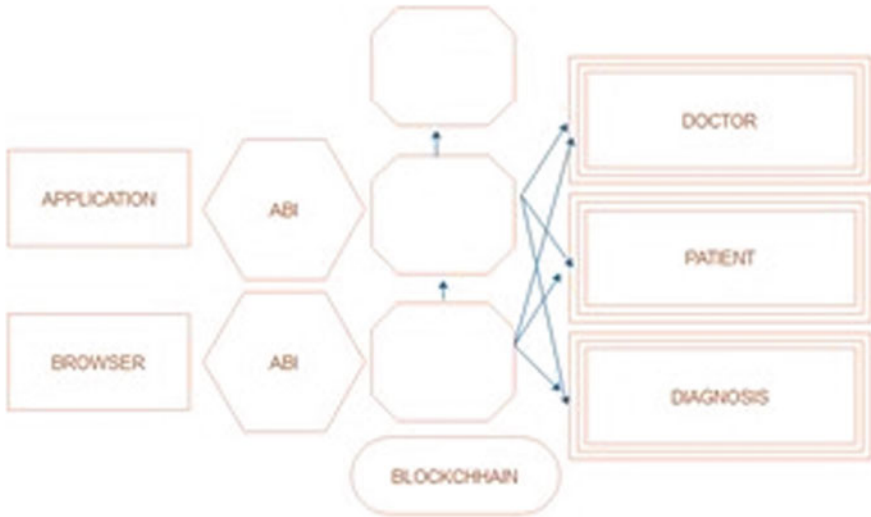
Fig. 3 Blockchain-based architecture of the system [13]

analyzed and designed by the monopolistic operator to meet EVs' individual energy demands while maximizing its utility. Figure 4 shows the system model of smart community.

D. *Discovery Process of Stock Value Information*

This paper [15] discusses the use of smart contract in stock value information. Currently, there are three steps for calculating stock value of any enterprise. It proposes use of blockchain in value record phase, as blockchain is distributed,
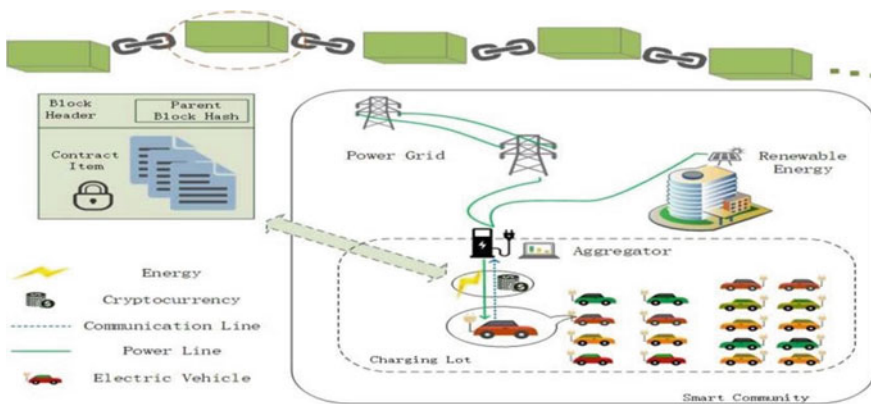
Fig. 4 System model of smart community [14]
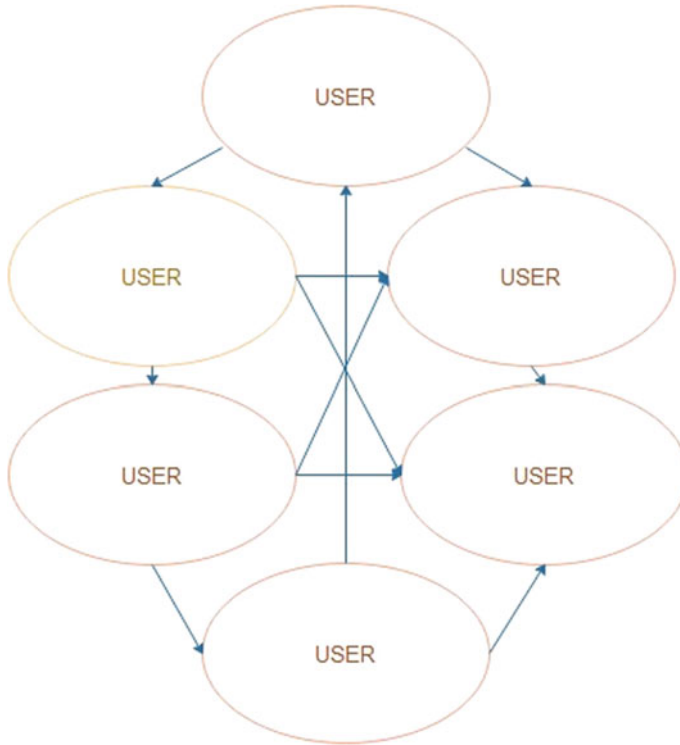
Fig. 5 Decentralized storage structure of blockchain [15]

immutable, and trustless environment where transaction information is stored in shared distributed ledger (Fig. 5).

4   Challenges in Implementing Smart Contract

1. Irreversibility: The very nature of blockchain transactions is irreversibility; if something has to be corrected in smart contract, then it has to be done via another transaction which also has to be stored in blockchain.
2. Digital Intelligence: Oracles as a service provide a platform which runs on the blockchain technology.
3. Legal Issues: Smart contract should be compatible with all legal policies and standards of their respective countries where they are executing.
4. Language of Smart Contract: Solidity is not flexible in terms of ontological expressiveness of business contracts and legal perspective of smart contract execution, so a more verifiable language is needed for implementing smart contracts.

5 Proposed Methodology

In this as shown in Fig. 6, we define some steps to follow a legally proof method for implementing smart contracts.

1. Select permission-based blockchain for the application (required).
2. Smart contract must be written by third party that must interpret the contract according with the intentions of the involving parties and legal policies both.
3. If the contracts conditions were determined by a neutral blockchain, then go to Step 4, otherwise go to Step 2.
4. If the agreement enforced through the terms and mechanism sets forth in terms of the contract itself and not by public law enforcers, then go to Step 2, otherwise go to Step 5.
5. Formation of the contract, clarity of intentions, and finality of intentions must be implemented.
6. Check and measure the performance of smart contract.
7. Smart contract is not necessarily perfect and must have flexibility to modification; if some damage occurs after executing the smart contract, then go to Step 2, otherwise go to Step 8.
8. Remedy judged by both the parties and legal policymakers must be implemented and executed.

Distributed Ledger(Permissioned Blockchain)

Smart contract must be written by Third party

Check Contracts conditions were determined and executed by Neutral Blockchain

Check if the agreement is enforced through the terms and mechanism of contract itself and not by public law enforcers

Formation of Contract by means of mutuality of obligations between parties

Check for Ambiguity ,clarity of intentions has been built and finality of intentions implemented

Measure performance and then corresponding modification

if some damage occurs after the execution of contract

Remedy enforcement by both third party and legal policy makers

Fig. 6 Smart contract designing methodology

6 Conclusion and Future Scope

This paper discusses implementations of smart contract in various use cases like distributed energy transaction implementing a microservices system in Solidity shows that microservices can be employed using blockchain, but cost and complexity measures have not been analyzed. Discovery process of stock value information, in that real-time automatic discovery of stock value as transaction could be employed in blockchain network. Secondly proposed methodology or steps would make designing of smart contract more efficient and legally compliant and to certain level solve the problem of policy forming for blockchain smart contracts.

This methodology can be further extended in more details that it would be make flexible for specific blockchain applications, also the cost of implementing smart contracts may be reviewed on different parameters. Methodology could be further converted in to an algorithm, and simulation would be done to find out the verifiability of that algorithm.

References

1. Kumar B et al (2018) An overview of smart contract and use cases in blockchain technology. In: IEEE 9th ICCCNT, IISC, Bengaluru
2. Mohanta et al (2018) An overview of smart contract and use cases in blockchain technology. In: IEEE—ICCCNT, Bengaluru
3. https://coinmarketcap.com/all/views/all/
4. Szabo N (1997) Formalizing and securing relationships on public networks. First Monday 2.9
5. Bahga A, Madisetti VK (2016) Blockchain platform for industrial Internet of Things. J Softw Eng Appl 9(10):533
6. Ellul J, Pace GJ, Alkyl VM (2018) A virtual machine for smart contract blockchain connected internet of things. In: 9th IFIP international conference on new technologies, mobility and security (NTMS). IEEE
7. Dickerson T, Gazzillo P, Herlihy M, Koskinen E (2018) How to add concurrency to smart contracts. Bull Eur Assoc Theor Comput Sci 124:22–33
8. Tschorsch F, Scheuermann B (2016) Bitcoin and beyond: a technical survey on decentralized digital currencies. IEEE Commun Surv Tutorials 18(3):2084–2123
9. Dixit et al (2018) A self-aware contract for decentralized peer-to-peer (P2P) commerce. In: 2018 IEEE 3rd international workshops on foundations and applications of self systems
10. Jimenez et al (2018) Implementation of smart contracts using hybrid architectures with on and off–blockchain components. In: IEEE 8th international symposium on cloud and service computing (SC2)
11. Dwivedi et al (2018) A legally relevant socio-technical language development for smart contracts. In: IEEE 3rd international workshops on foundations and applications of self* systems
12. Yu S, Yang S, Li Y, Geng J (2018) In: Osgood Jr RM (ed) China international conference on electricity distribution. Tianjin, Berlin, Germany, 17–19 Sep 2018
13. Tonelli R et al (2019) Implementing a microservices system with blockchain smart contracts. In: IEEE IWBOSE 2019, Hangzhou, China
14. Yuntao et al (2018) Contract based energy blockchain for secure electric vehicles charging in smart community. In: IEEE 16th international conference on dependable, autonomic & secure computing

15. Liu X, Lin N, Wegmuller M (2018) An automatic discovery process of stock value information with software industry based on blockchain. IEEE
16. Kosba A et al (2016) Hawk: the blockchain model of cryptography and privacy-preserving smart contracts. In: IEEE symposium on security and privacy (SP). IEEE

Deep Learning Neural Networks for Detection of Onset Leukemia from Microscopic Images

K. Sujatha, N. P. G. Bhavani, V. Srividhya, B. Latha, M. Sujitha, U. Jayalatsumi, T. Kavitha, A. Ganesan, A. Kalaivani, B. Rengammal Sankari, and Su-Qun Cao

Abstract The major challenge in the medical field is the detection of blood cancer or leukemia in its early stage detection. Once the blood cancer has been diagnosed in its early stage, treatment becomes easy and also it works out well. Presently, only manual identification of blood cancers from microscopic images is available. Manual identification, at times many lead to improper diagnosis. The automation of blood cancer identification depends on the availability of the database to form a medical expert system. Leukemia is caused due to deficiency of white blood cells in the blood stream in human body. This research work focuses on detecting the

K. Sujatha (✉) · S.-Q. Cao
Department of Electrical and Electronics Engineering, Dr. M.G.R. Educational and Research Institute, Chennai, Tamilnadu 600095, India
e-mail: sujathak73586@gmail.com

N. P. G. Bhavani
Department of ECE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

V. Srividhya · B. Latha
Department of Electrical and Electronics Engineering, Meenakshi College of Engineering, Chennai, India

M. Sujitha
Department of Physics/ECE, Dr. M.G.R. Educational and Research Institute, Chennai, Tamilnadu 600095, India

U. Jayalatsumi
Department of ECE, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India

T. Kavitha
Department of Civil Engineering, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India

A. Ganesan
Department of EEE, RRASE College of Engineering, Chennai, Tamil Nadu, India

A. Kalaivani
Department of CSE, Rajalakshmi Institute of Technology, Chennai, India

B. Rengammal Sankari
Faculty of Electronic Information Engineering, Huaiyin Institute of Technology, Huaiyin, China

deficiency of white blood cells at the initial stage itself. For this reason, fuzzy C-means (FCM) algorithm and active contours are used. The features are extracted from the region of interest (RoI) by segmentation of the microscopic images. The outcome of the proposed visual recognition system is that the features are extracted from the specific region of the microscopic blood cell image using FCM algorithm to detect the blood cancer at the early stage using Convolutional Neural Networks (CNN). Various algorithms are implemented and a trail has been conducted using standard simulation package to demonstrate that active contours yield optimal results with high accuracy, recall, and a low rate of false classification ratio.

1   Introduction

The microscopic examination of blood sample offers clinical information for diagnosing the health conditions of the concerned patient. The investigative results from blood sample analysis give a detailed report of haematic pathologies under critical conditions. The count of the leucocytes in the blood sample indicates the various stages of leukemia which is diagnosed by analyzing the classification's algorithm output. The expert medical system will be capable to handle a large database as compared with the conventional differential method of blood cell count. A small area, the region of interest (RoI) is identified for counting the leucocytes in the blood smear. This counting of the leucocytes is a manual process and may cause human errors and optical system errors. The error free counting of the blood cells depends on the efficiency of the lab technicians involved in the sampling process. As a result, there is a need for developing an indigenous automated differential counting system. Blood cancer is a dreadful disease that is prevalent among the people and left untreated in its initial stage leads to severe complications in children as well as in adults. In nature, the old leucocytes present in the blood stream are worn out and eliminated followed by the formation new leucocytes. The condition where the old white blood cells (WBCs) or leucocytes when worn out, remain in the blood stream and not being replaced by the new ones, reduces the prescribed count of the WBCs in the blood stream. This kind of growth does not provide space for the newly formed leucocytes in the cytoplasm. The newly formed WBCs will be multiplying at a very faster rate without the old ones being destroyed. This causes leukemia or blood cancer.

2 Literature Review

An indicative result from radiology dispenses the various images corresponding to the medical images to offer specific diagnosis. The researcher, Jagadeesh et al. [1], has provided some strategies for noise removal from the images so as to detect the leukemia with required description. The methods followed include segmentation [1] using watershed algorithm to identify the geometric patterns of the normal and abnormal WBCs. The image processing techniques include image illustration, enrichment, segmentation, and many other morphological operations. The acquired MRI or CT images are analyzed using the above mentioned image processing techniques to detect the abnormalities like tumor detection, arteries block, and fractured bones. Parameshwarappa and Nandish [2] proposed a method using Fourier transform to identify the abnormalities in the brain. Patil and Jain [3] proposed a technology to determine the abnormal growth in lungs so that an appropriate treatment can be offered. For determination of lung cancers at early stages, some segmentation and thresholding algorithms are used. Similarly, another researcher has implemented a mathematical approach to detect the edges of the tumors in lung from MRI or CT images [4]. The salt and pepper noise which is most prevalent in medical images is removed by using effective filtering techniques [4]. Morphological methods are used to detect the edges in case of cancerous cells. Morphological degeneration is a common drawback associated with salt and pepper noise. Hence, image restoration followed by de-noising and edge detection can evolve some morphological algorithms to detect the cancerous cells [5, 6].

3 Algorithms

3.1 Fuzzy C-Means Algorithm

Step 1: Assign K value and membership function and identify the center for the cluster.

$$J = \sum_{(i=1)}^{N} \sum_{(j=1)}^{C} \left\| x_j - c_j \right\|^2 \tag{1}$$

N	denotes the data
C	denotes the clusters
δ_{ij}	Membership value
$\left\| x_j - c_j \right\|$	denotes the closeness of the cluster 'i' with the cluster 'j'.

Step 2: Calculate the cluster center with maximum membership function.

$$\delta_{ij} = 1/\sum_{(k=1)}^{C} \|(x(i) - C_i)/(x(i) - C_k)\|^{(2/(m-1))} \qquad (2)$$

m is the fuzziness coefficient and c_j is calculated as follows:

$$C_j = \left(\sum_{(i=1)}^{N} (\delta_{ij}^m [\![x_j]\!])\right) / \left(\sum_{(i=1)}^{N} \delta_{ij}^m\right) \qquad (3)$$

δ_{ij}^m denotes the degree of membership value and is between 0 and 1.

$$\sum \delta_{ij} = 1$$

Step 3: Coefficients are such that $1 < m < \infty$.

$$\varepsilon = \Delta i^{(N)} \Delta j^c \left|\delta_{ij}^{(k+1)} - \delta_{ij}^k\right| \qquad (4)$$

$\delta_{ij}^{(k+1)}, \delta_{ij}^k$ membership functions for $K, K + 1$ iterations.

Step 4: When the cluster center is stabilized, stop the algorithm.

3.2 Active Contours

The parameters are fine tuned for appropriateness using the following steps.

Step 1: Keep the snake at the region of interest (RoI).

Step 2: Move the snake on iterative basis using the forces.

Step 3: Find the suitable energy function which accurately corresponds to the ROI.

Step 4: Continue till the energy function ceases and is close to the ROI.

3.3 Convolutional Neural Network (CNN)

Compute the net value

$$y_j(w, x, a_j) = f\left(\sum_{(i=1)}^{(r^2)} (w_{ji}x_i) + a_j\right) \qquad (5)$$

w_{ji}	the weight between the pixel and hidden neuron, j, i of the input image.
x_i	gray value of the input pixel i.
a_j	the bias of the hidden neuron j.
$x1, x2, \ldots x_{(r^2)}$	Input image pixels, each connected with every neuron j.

The output layer is completely associated with the hidden layer. The sigmoid activation function, z_0, of the output neuron is characterized by

$$z_0(w, y, g_0) = 1 / \left(1 + \exp\left\{ -\left[\sum_{(j=1)}^{(nN^2)} (w(0_j)y_j) + g0 \right] \right\} \right) \qquad (6)$$

$w(0_j)$	Weight between the neuron and the output neuron in the hidden layer j
nN^2	total number of neurons in the hidden layer
g_0	bias of the output neuron.

The optimization goal is the Mean Squared Error (MSE).

4   Methodology

As a result, designed system incorporates training and testing. At the beginning, microscopic images of the blood are subjected to preprocessing for noise removal. Following this, color images are converted into gray scale images. Features are extracted and reduced using active contour modeling. These features are then used to train and test the CNN classifier to detect the normal, acute, and chronic blood cancer. Figure 1 describes the architecture of proposed system. Acute leukemia is one in which the WBCs are dividing and spreading the disease at a faster rate. The reference count of WBCs is given in Table 1. The major symptom is that the patient will be suffering from frequent illness. Chronic leukemia has WBCs which contain both immature and mature cells. For years together, the chronic leukemia will not show any symptoms. It is prevalent in adults rather than in children.

5   Results and Discussion

The microscopic image of the blood is depicted in Fig. 1 which shows the microscopic input image. In the proposed method, microscopic image is enhanced using preprocessing steps (Fig. 2), they are color conversion, filtering, and histogram equalization. Followed by segmentation using fuzzy C-means algorithm, Fig. 3 shows the enhancement done to improve contrast. For using fuzzy C-means algorithm, three classes are selected. Figures 4, 5, and 6 show using fuzzy C-means segmentation
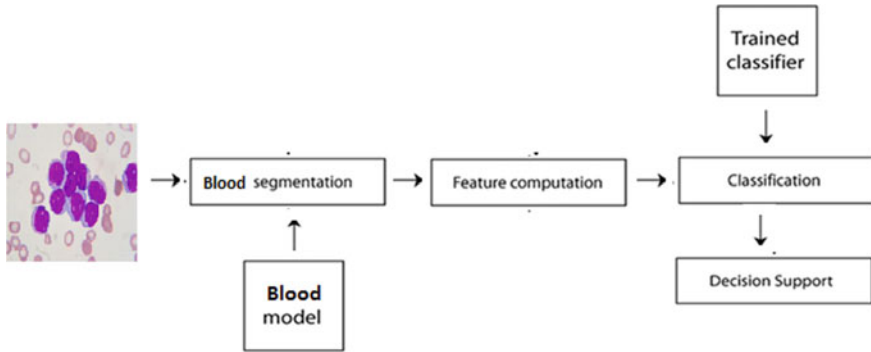
Fig. 1 Schematic for blood cancer detection using image processing

Table 1 Reference count for leucocytes

Approximate low range	< 4000 white blood cells per mm^3
Approximate normal range	4500–10,000 white blood cells per mm^3
Approximate high range	> 10,000 white blood cells per mm^3
Normal range	3700–10,500

outputs. Here, C is taken as 3 because to get the proper region of interest, in this case it is nucleus. After obtaining segmentation, Fig. 4 is taken for feature extraction as it has complete nucleus which is the region of interest. Feature extraction is done and these features (Table 2) are used as inputs to classify input microscopic image using CNN as normal, acute, and chronic stages of blood cancer.

Fig. 2 Input image
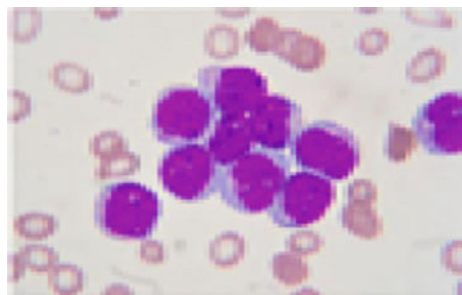
Fig. 3 Output for histogram equalization

Fig. 4 RoI—normal and abnormal images
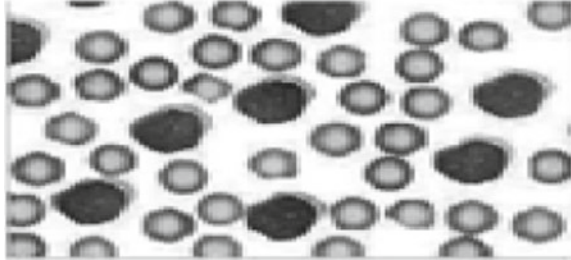
Fig. 5 Fuzzy clustering
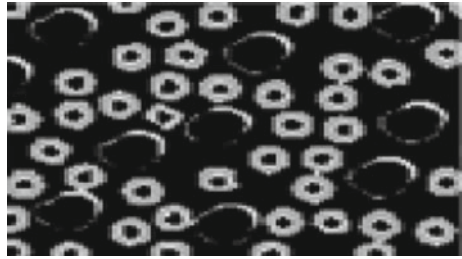
Fig. 6 Original image

Table 2 Features of normal and abnormal images

Images	Features for normal images of blood cells				Features for abnormal images of blood cells (acute stage)			
	Mean	SD	PSNR	MSE	Mean	SD	PSNR	MSE
IMG 1	130.25	52.69	2.78	190	146.68	66.98	77	0.34
IMG 2	131.26	50.80	3.90	189	143.46	62.67	50	0.37
IMG 3	121.625	45.52	3.44	182	149.23	62.39	78	0.45
IMG 4	123.548	56.62	3.89	178	141.72	63.25	89	0.56
IMG 5	112.761	46.89	4.01	158	142.25	60.35	91	0.44
Images	Features for abnormal images of blood cells (chronic stage)			
IMG 1	Mean	SD	PSNR	MSE
IMG 2	176.88	46.98	121	0.74
IMG 3	173.48	52.67	125	0.77
IMG 4	179.22	42.39	127	0.75
IMG 5	171.72	43.25	129	0.76

Fig. 7 Segmentation using active contour
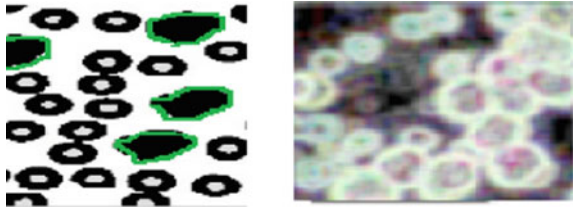
6 Conclusion

Hence, the proposed work uses active contour and fuzzy C-means clustering algorithms. Classified results from CNN provide high accuracy, improved peak signal to noise ratio (PSNR), and optimal value of MSE. The proposed work provides 97.76% of accuracy (Fig. 7).

References

1. Jagadeesh S, Nagabhooshanam E, Venkatchalam S (2013) Image processing based approach to cancer cell prediction in blood sample. Int J Technol Eng Sci 1(1). ISSN: 2320-8007
2. Parameshwarappa V, Nandish S (2014) A segmented morphological approach to detect tumour in brain images. Int J Adv Res Comput Sci Softw Eng 4(1). ISSN 2277 128X
3. Patil BG, Jain SN (2014) Cancer cells detection using digital image processing methods. Int J Latest Trends Eng Technol (IJLTET) 3. ISSN: 2278-621X
4. Sivappriya T, Muthukumaran K (2014) Cancer cell detection using mathematical morphology. Int J Innov Res Comput Commun Eng 2(1)

5. Chandhok C, Chaturvedi S, Khurshid AA (2012) An approach to image segmentation using k-means clustering algorithm. Int J Inf Technol (IJIT) 1(1)
6. Dandotiya R, Singh S, Jalori S (2014) A survey of image classification techniques for flood monitoring system. International conference on emerging trends in computer and image processing

Design of 4 × 4 Butler Matrix for 5G High Band of 26 GHz

P. Manjunath, Dhruv Sharma, and P. Shanthi

Abstract A 4 × 4 Butler matrix for high band 5G frequency of 26 GHz has been designed and simulated using ADS tool with substrate material Rogers RO4300 ($\varepsilon_r = 3.38$). The subcomponents of Butler matrix, i.e., branch line coupler (BLC), phase shifter ($- 45°$) and 0 dB crossover were designed individually and integrated to get the phase differences of $- 135°$, $- 45°$, $45°$ and $135°$ at the output ports with respect to excitation of input ports. Rectangular patch antenna was also designed in ADS tool for 26 GHz with same substrate and was integrated with Butler matrix to get the desired beam patterns according to the excitation.

Keywords Butler matrix · Patch antenna · Beam patterns · Coupler · Phase shifter · Crossover

1 Introduction

Improvement of communication in the undesirable conditions is becoming one of the important tasks as the communication systems are increasing day by day. One of the best techniques that can be adopted for solving this problem is smart antennas. The smart antennas can be broadly classified into adaptive arrays and switched-beam systems. Adaptive antenna arrays perform better by rejecting the interference and adapting according to the environment. But this requires huge signal processing and is costly. On contrast, a switched-beam array produces selected beams directed according to requirement and can be switched easily. But these are not as efficient as the adaptive arrays.

P. Manjunath (✉) · D. Sharma · P. Shanthi
Department of Electronics and Telecommunication Engineering, RV College of Engineering, Bengaluru, India
e-mail: pmanjunath.lrf20@rvce.edu.in

D. Sharma
e-mail: dhruvshrama.lrf20@rvce.edu.in

P. Shanthi
e-mail: shanthip@rvce.edu.in

Butler matrix is one of the well-known beamforming networks for use in switched-beam antenna systems. An $N \times N$ Butler matrix consists of N input ports and N output ports feeding N antennas where all paths between input and output ports are equal. The Butler matrix may be made using planar or waveguide technology. Using planar structure, the Butler matrix is composed by 3 main elements: 3 dB/90° quadrature couplers, crossovers and phase shifters. In order to achieve wideband features, both wideband 3 dB couplers and crossovers are required [1]. It can be used to generate N beams to improve carrier-to-interface ratio and frequency reuse in cellular wireless communication systems [2].

The dimensions of the matrix are determined by the sizes of all directional couplers included in its composition, and the lower the frequency, the larger the area that it will occupy on the microwave substrate. Therefore, improving their design, miniaturization and increasing technological feasibility are an urgent task for microwave technology [3].

In Fig. 1, the architecture of the Butler matrix is shown. The 4×4 Butler matrix can be used at transmitter or receiver and can also be used symmetrically; i.e., the input and output ports can be used interchangeably. Figure 1 shows that four branch line couplers, two delay line phase shifters and two crossovers have been used in the design of this Butler matrix. The rectangular patch antennae were designed after the verification of progressive phase shifts between the output ports.

The switched-beam arrays are proposed to eliminate the limitation of huge processing, and by judging power levels of electromagnetic waves, the main beam of the radiation pattern of beam switched antenna can point to the optimal direction and then build up the communication link [4]. Butler matrices can also be designed with variations in the subcomponents as Schiffman phase shifters were used in [5] to improve the efficiency. The phase shifter subcomponent can be varied and with
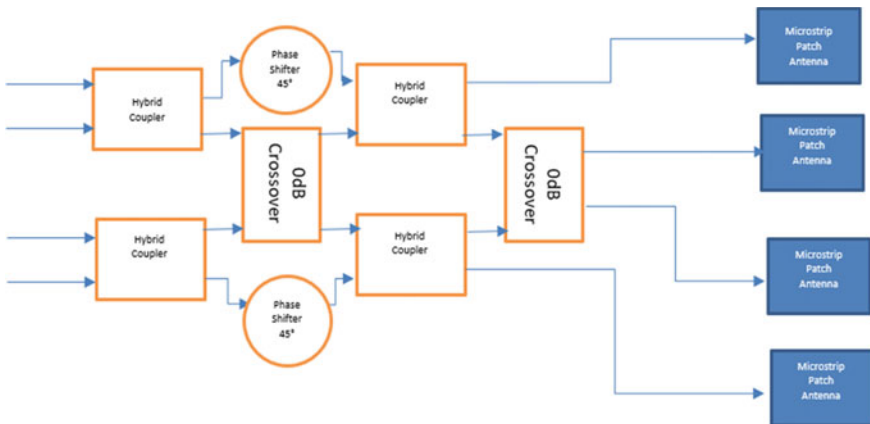
Fig. 1 Architecture of 4×4 butler matrix with input ports (at left) and output ports (at right) connected to antennas

that in our control beam patterns can be modified accordingly. This was shown in [6] with an asymmetrical Butler matrix design.

Butler matrices come under the technique of space division multiple access (SDMA). The SDMA technique allows discrimination of the users standing at different angular positions around the base station and makes possible the reuse of the existing bandwidth toward them. As a result, system's capacity is expanded [7]. The implementation of Butler matrices in the handheld devices for 5G communication is becoming very necessary and [8] shows the technique of implementing using substrate integrated technology. A similar approach is seen [9] where a 3 × 3 matrix is designed for 5G mobile devices. As requirements of 5G systems are growing, the Butler matrices are designed for 5G high band of 28 GHz in [10, 11]. The other novel technique of improving the efficiency is stack the components in layers which was presented in [12] with good efficiency.

2 Design of Butler Matrix and Its Subcomponents

The first step toward the designing is to consider a substrate material. In this case, Rogers RO4300 has been considered which has an $\varepsilon_r = 3.38$. The substrate was necessarily selected to perform better at high frequencies as our requirement was to perform at 26 GHz. The thickness of the substrate is taken as 0.203 mm. With these considerations, design of each subcomponent is shown further.

2.1 Design of Branch Line Coupler

The branch line coupler schematic for the desired values is calculated. The return loss expected was less than − 30 dB, and coupling and through were expected to be − 3 dB as the necessity was to get half power with 90° phase shift. The isolation is expected to be less than − 40 dB. The length and width calculations for expected results are tabulated in Table 1. The S-parameter matrix obtained after even–odd mode analysis of branch line coupler is given in Eq. (1).

$$S = \frac{-1}{\sqrt{2}} \begin{bmatrix} 0 & j & 1 & 0 \\ j & 0 & 0 & 1 \\ 1 & 0 & 0 & j \\ 0 & 1 & j & 0 \end{bmatrix} \tag{1}$$

The horizontal lines and four port lines are of impedance shown in Eq. (2)

$$Z_0 = 50 \, \Omega \tag{2}$$

Table 1 Specification of butler matrix subcomponents

S. No.	Parameters	Values
1	Operating frequency	26 GHz
2	Substrate specification	Rogers RO4300 Dielectric constant = 3.38 Thickness = 0.203 mm
3	Conductor	Copper Thickness = 35 μ, Conductivity = 5.8×10^7
4	BLC measurements	$Z_0 = 50\ \Omega$ $W = 0.440428$ mm $L = 2.01924$ mm $Z_1 = 35.35\Omega$ $W = 0.769475$ mm $L = 1.06066$ mm
5	Crossover measurements	$\lambda/4$ line $W = 0.440528$ mm, $L = 1.57742$ mm
6	Phase shifter measurements	$W = 0.440528$ mm
7	Patch antenna measurements	$W = 3.896$ mm $L = 3.068$ mm

and the $\lambda/4$ lines are of impedance as shown in Eq. (3)

$$Z_1 = Z_0/\sqrt{2} = 35.35\ \Omega \tag{3}$$

The width and length are calculated using Eqs. (4–6)

$$\left(\frac{w}{d}\right) = \frac{8e^A}{e^{2A} - 2} \tag{4}$$

where

$$A = \frac{Z_0}{60}\sqrt{\frac{\varepsilon_r + 1}{2}} + \frac{\varepsilon_r + 1}{\varepsilon_r - 1}\left(\frac{0.11}{\varepsilon_r} + 0.23\right)$$

Z_0 is selected accordingly.

$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(\frac{1}{\sqrt{1 + \frac{12d}{w}}}\right) \tag{5}$$

The width w for (5) is obtained from (4). Further, λ_g is calculated using this effective permittivity obtained in (5) using (6) as follows

$$\lambda_g = \frac{c}{f\sqrt{E_{\text{eff}}}} \tag{6}$$

The schematic was obtained with these design considerations, and it was obtained with the help of line calculator tool in ADS with substrate and frequency defined as per the application. The schematic of branch line coupler designed for 26 GHz is as shown in Fig. 2.

The layout conversion was obtained by the automatic conversion of ADS tool, and the layout obtained is shown in Fig. 3.
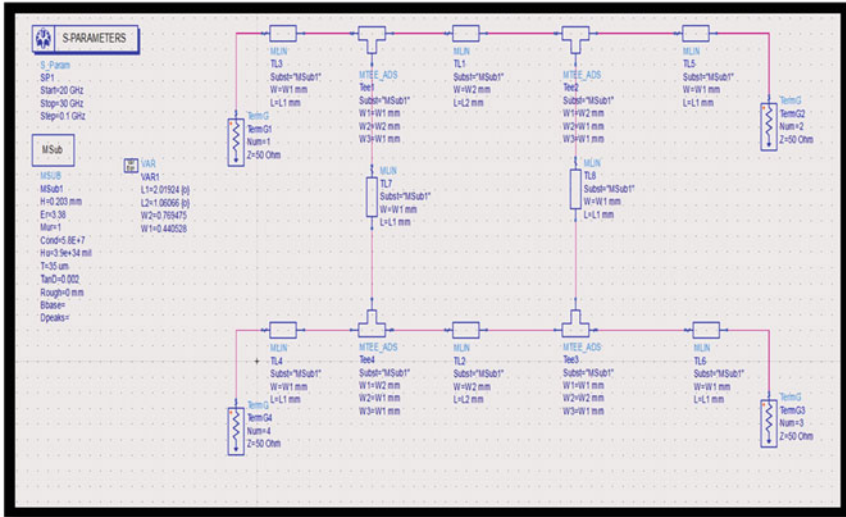
Fig. 2 Schematic of the branch line coupler
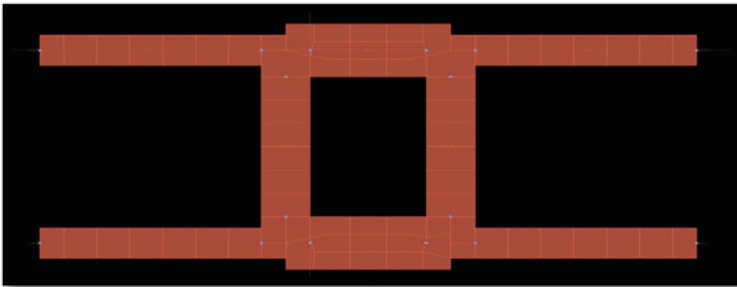
Fig. 3 Layout of the branch line coupler

2.2　Design of 0 dB Crossover

The crossover schematic for the desired values is calculated. The return loss expected was less than − 30 dB, and coupling and through were expected to be − 10 dB and 0 dB. The isolation is expected to be less than − 25 dB. The length and width calculations for expected results are tabulated in Table 1. The *S*-parameter matrix of 0 dB crossover is given in Eq. (7).

$$S = \begin{bmatrix} 0 & 0 & j & 0 \\ 0 & 0 & 0 & j \\ j & 0 & 0 & 0 \\ 0 & j & 0 & 0 \end{bmatrix} \tag{7}$$

The λ/4 line impedances are calculated in a similar way as that of BLC as 0 dB crossover is joining of two BLCs. The width and length are calculated using Eqs. (4–6).

The schematic was obtained with these design considerations, and it was obtained with the help of line calculator tool in ADS with substrate and frequency defined as per the application. The 0 dB crossover designed for 26 GHz is as shown in Fig. 4.

The layout conversion was obtained by the automatic conversion of ADS tool, and the layout obtained is shown in Fig. 5.
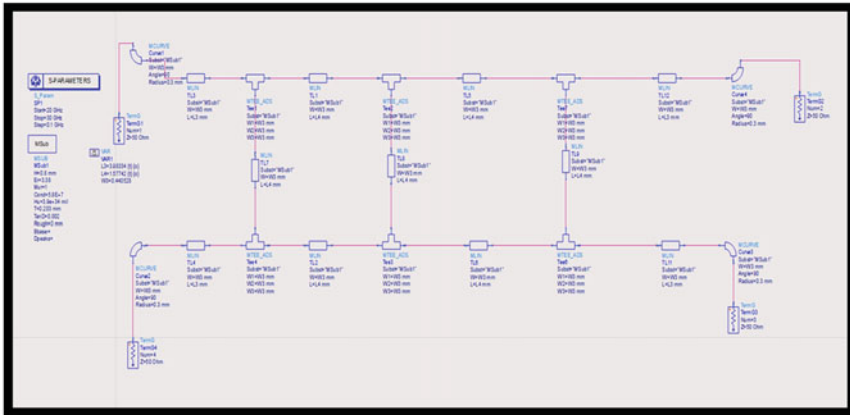
Fig. 4　Schematic of the 0 dB crossover
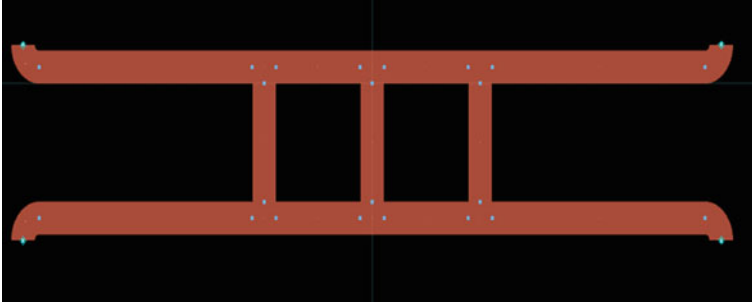
Fig. 5 Layout of the 0 dB crossover

2.3 Design of Phase Shifter

The phase shifter schematic for the desired values is calculated. The phase shift S_{21} expected was $-45°$. The length and width calculations for expected results are tabulated in Table 1. The phase shifter is implemented using Eq. (8) which is as follows.

$$\varphi = \frac{2\pi}{\lambda} L \qquad (8)$$

where

$$\lambda = \frac{\lambda_0}{\sqrt{\varepsilon_{\text{eff}}}}$$

and ε_{eff} is obtained from Eq. (5).

The length obtained from Eq. (8) is used for designing the delay line with required bend for symmetry. This can be made by adding additional line without changing the original requirement. The schematic was obtained with these design considerations, and it was obtained with the help of line calculator tool in ADS with substrate and frequency defined as per the application. The 45° phase shifter designed for 26 GHz is as shown in Fig. 6.

The layout conversion was obtained by the automatic conversion of ADS tool, and the layout obtained is shown in Fig. 7.

2.4 Design of Rectangular Patch Antenna

The tabulated values of expected requirements for the patch antenna are shown in Table 1. The return loss expected is less than -10 db. The layout was constructed based on the theoretical calculation done using Eqs. (9–13). Coordinates obtained
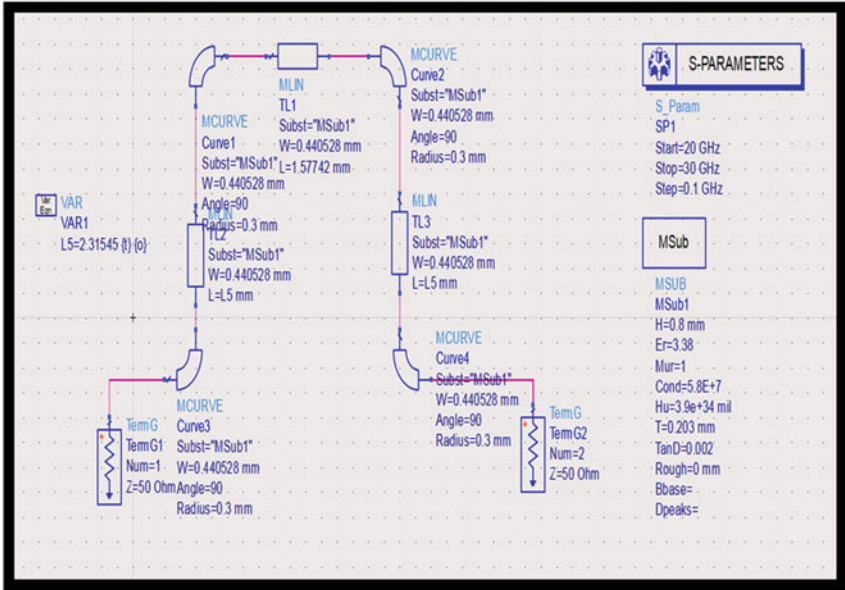
Fig. 6 Schematic of the phase shifter

Fig. 7 Layout of the phase shifter
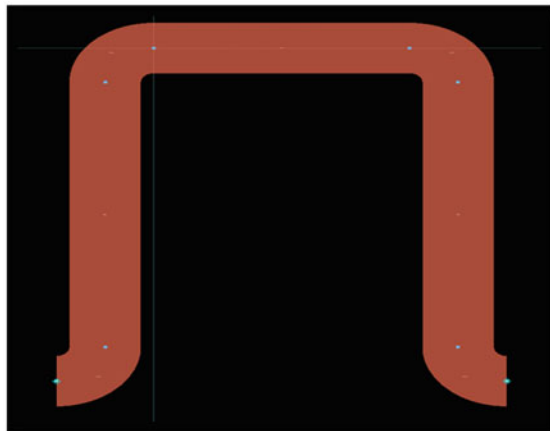
were calculated based on the measurements obtained, and the layout obtained is shown in Fig. 8.

$$f_0 = 0.3\sqrt{\frac{Z_0}{h\sqrt{\varepsilon_r - 1}}} \qquad (9)$$

Fig. 8 Layout of the rectangular patch antenna
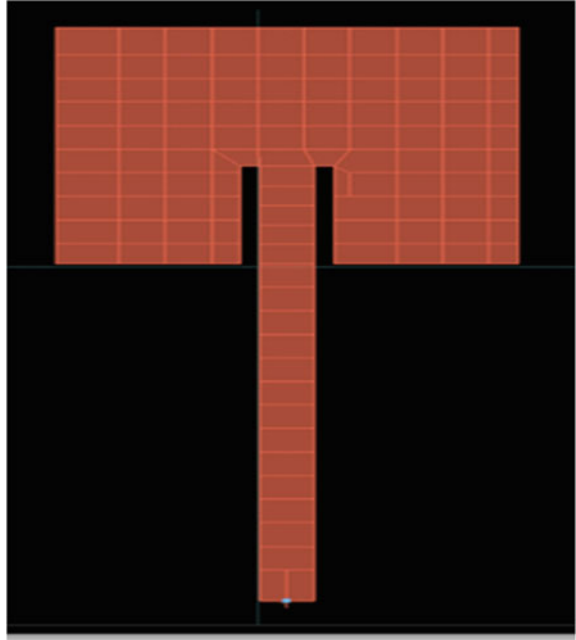
where $f_0 = 26$ GHz

Width is obtained by Eq. (10)

$$W = \frac{V_0}{2 f_r} \sqrt{\frac{2}{\varepsilon_r + 1}} \qquad (10)$$

where 'V_0' is the free space velocity.

Extended incremental length is obtained from Eq. (11)

$$\frac{\Delta l}{h} = 0.421 \left(\frac{+0.3}{\varepsilon_{\text{eff}} - 0.258} \right) \left(\frac{\frac{w}{d} + 0.264}{\frac{w}{d} + 0.818} \right) \qquad (11)$$

The actual length is obtained from Eq. (12)

$$L = \frac{1}{2 f_r \sqrt{\varepsilon_{\text{eff}}} \sqrt{\mu_0 \varepsilon_0}} - 2 \qquad (12)$$

The effective length is given by Eq. (13)

$$L_e = L + 2\Delta l \qquad (13)$$

2.5   Design of Overall Butler Matrix

The complete Butler matrix is obtained by integrating the subcomponents in the desired fashion as mentioned in the architecture. This was done as schematic initially in ADS tool with the integration of all the subcircuits designed above. The final version was optimized to get the desired output. The complete schematic is as shown in Fig. 9.

The layout conversion was obtained by the automatic conversion of ADS tool, and the layout obtained is shown in Fig. 10.
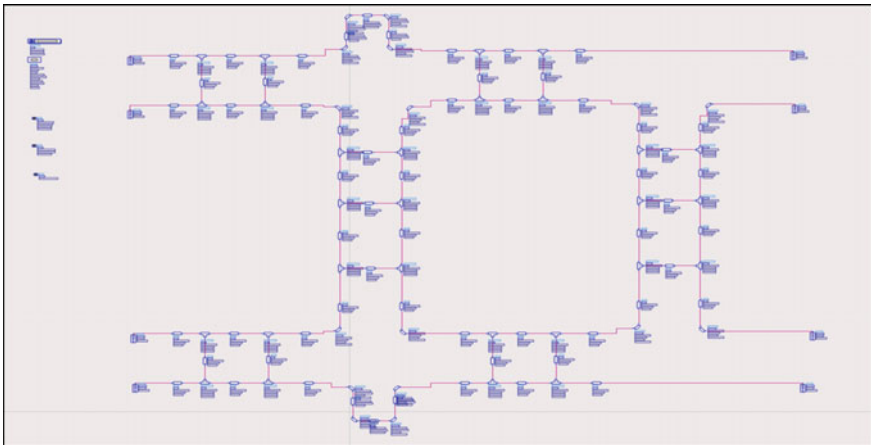
Fig. 9   Schematic of the overall butler matrix
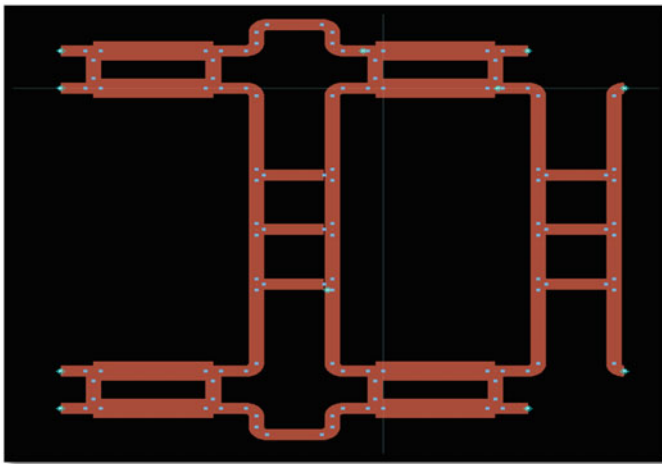
Fig. 10   Layout of the overall butler matrix
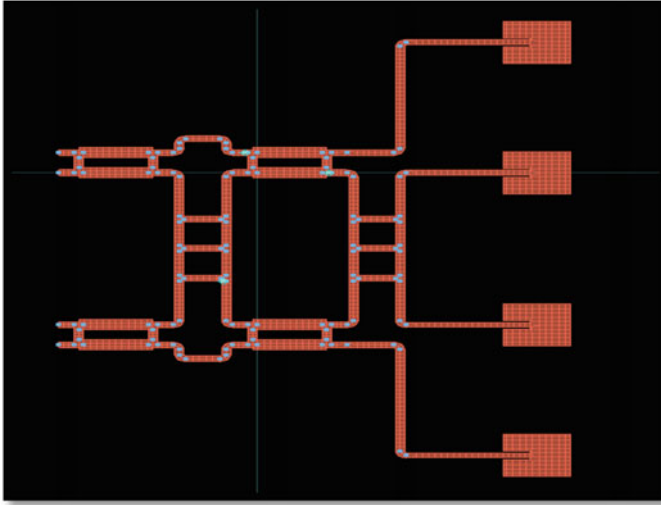
Fig. 11 Final layout after the integration of patch antennas

Further, the verification of phase differences between the ports was done and the patch antennas were integrated in the layout. The lengths were added without any effect to the phase differences between the ports so that the antennas are equidistant and antenna array giving broadside emission. The final layout was obtained as shown in Fig. 11.

The radiation pattern can be obtained by knowing the radiation pattern of single element and multiplying with the array factor. The array factor equation for four elements is shown on Eq. (14).

$$AF = \sum_{n=1}^{4} e^{j(n-1)\varphi} \tag{14}$$

where $\varphi = \mathrm{kd}\cos\theta + \beta$ and $k = 2\pi/\lambda$, d is the distance between the antennas, and β is the progressive phase shift which is $-135°, -45°, 45°, 1$ and $35°$ in this case.

3 Results and Analysis

The results of each subcircuit and the overall Butler matrix are discussed as follows. The design of branch line coupler for 26 GHz is done based on the expected results as mentioned in the design. The results were optimized in ADS tool after the manual calculation to meet the exact requirements. The results of BLC are as shown in Fig. 12. Obtained results are tabulated in Table 2.
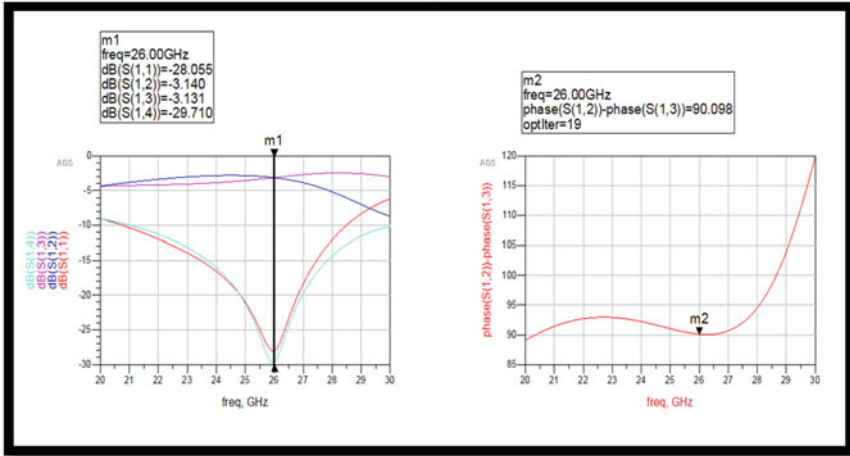
Fig. 12 Results of branch line coupler

Table 2 Subcomponent and overall butler matrix results

S. No.	Subcircuit	Results
1	Branch line coupler	$S11 = -28.055$ dB $S12 = -3.140$ dB $S13 = -3.131$ $S14 = -29.710$ dB
2	0 dB crossover	$S11 = -25.042$ dB $S12 = -12.637$ dB $S13 = -0.515$ dB $S14 = -20.487$ dB
3	Phase shifter	Phase $(S_{12}) = -45.855°$
4	Butler matrix	$S16$-$S15 = 161.7°$ $S17$-$S16 = 99.3°$ $S18$-$S17 = 84.504°$ $S26$-$S26 = -45.831°$ $S27$-$S25 = -33.639°$ $S28$-$S27 = -44.260°$ $S36$-$S35 = 44.260°$ $S37$-$S36 = 33.639°$ $S38$-$S37 = 45.831°$ $S46$-$S45 = -84.504°$ $S47$-$S46 = -161.790°$ $S48$-$S47 = -99.300°$
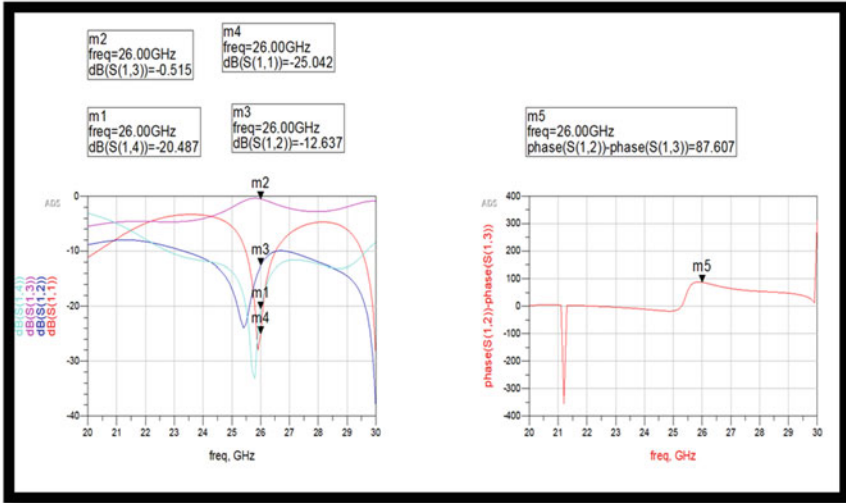
Fig. 13 Results of 0 dB crossover

0 dB crossover was designed as per the requirements, and the final result was optimized in ADS tool. The obtained results are shown in Fig. 13 and are tabulated in Table 2.

Delay line phase shifter is designed to get the angle of S_{21} as $-45°$. The obtained phase result is as shown in Fig. 14, and the result is tabulated in Table 2.

These three subcomponents were integrated, and each port was excited one after the other to check the phase differences between each output port. The obtained phase differences are progressive phase shifts as shown in Fig. 15 and are tabulated in Table 2.
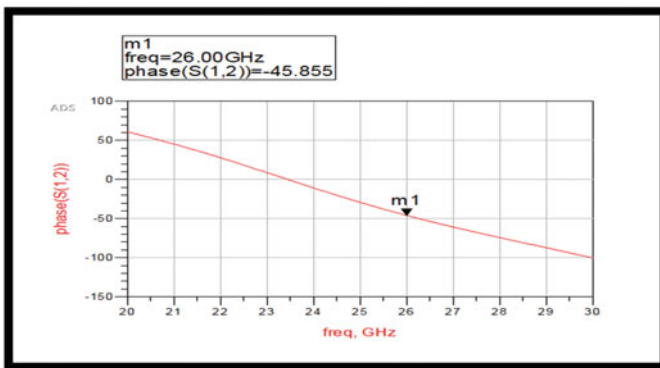
Fig. 14 Results of phase shifter
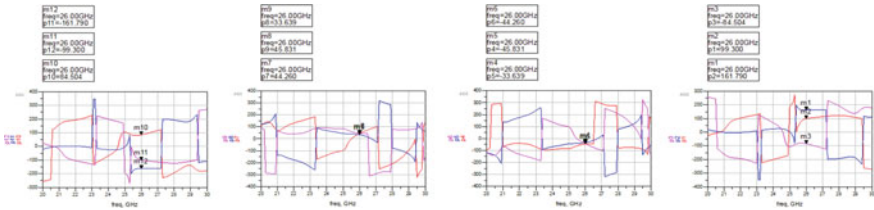
Fig. 15 Progressive phase shifts with respect to port of excitation
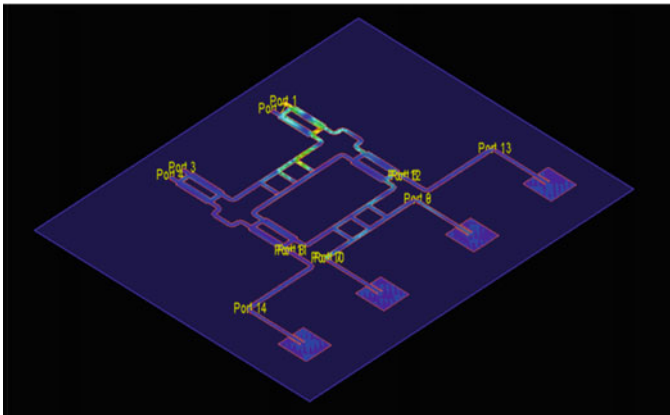
Fig. 16 Field distribution when port 1 is excited

The rectangular patch antennae designed for 26 GHz were integrated with the Butler matrix, and the magnitudes were observed as same at all the antennae. The phases were different according to the port excited but a progressive phase shift was observed which is necessary for broadside emission. The field distribution diagram when port 1 is excited as an example is shown in Fig. 16.

The radiation pattern in 3D was obtained in ADS tool with each port excited. 2D cuts were made in the desired theta and phi direction to get the 2D radiation plots as shown in Fig. 17. Few sidelobes were observed when ports 2 and 3 were excited along with the reduction in magnitudes.

4   Conclusion

A beam shifting 4×4 Butler matrix was implemented that manifests proper phase shifts at excitation of every port as seen in the results. The designed Butler matrix consists of a phase shifter, 0 dB crossover, branch line coupler and microstrip antennae which are all integrated together to yield the appropriate results. The
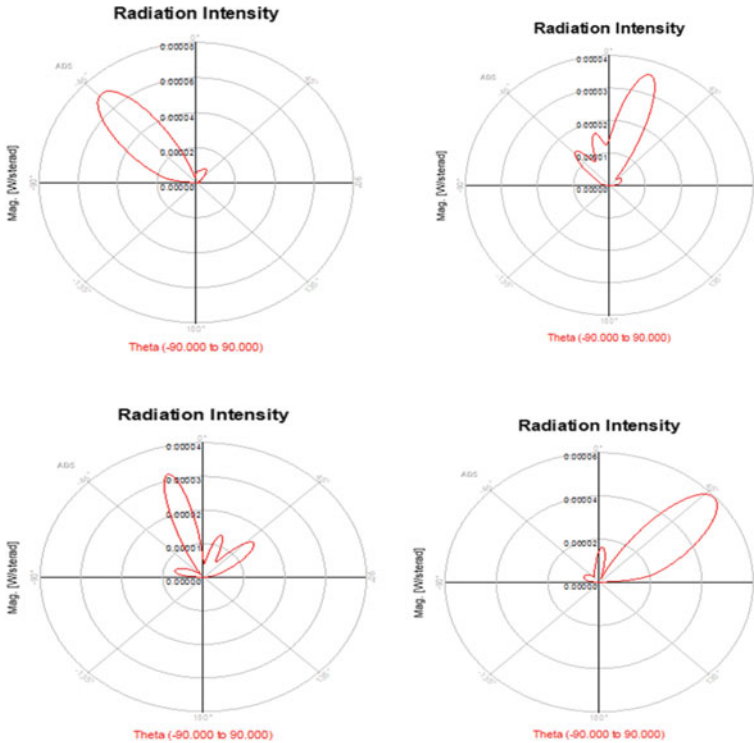
Fig. 17 Radiation pattern when ports 1, 2, 3 and 4 are excited, respectively

obtained radiation patterns show an accurate and efficient phase shift as there are only small or reduced sidelobes present which also inculcates lesser power wastage.

References

1. Nachouane H, Najid A, Tribak A, Riouch F (2014) Broadband 4×4 Butler matrix using wideband 90° hybrid couplers and crossovers for beamforming networks. Int Conf Multimedia Comput Syst (ICMCS) 2014:1444–1448. https://doi.org/10.1109/ICMCS.2014.6911309
2. Zulfi AD, Munir A (2020) Implementation of meander line structure for size miniaturization of 4×4 butler matrix. In: 2020 27th international conference on telecommunications (ICT), pp 1–4. https://doi.org/10.1109/ICT49546.2020.9239448
3. Ravshanov DC, Letavin DA, Terebov IA (2020)Butler matrix 4x4 ultra high frequency. In: 2020 systems of signal synchronization, generating and processing in telecommunications (SYNCHROINFO), pp 1–4. https://doi.org/10.1109/SYNCHROINFO49631.2020.9166084
4. Huang F, Chen W, Rao M (2016) Switched-beam antenna array based on butler matrix for 5G wireless communication. In: 2016 IEEE international workshop on electromagnetics: applications and student innovation competition (iWEM), pp 1–3. https://doi.org/10.1109/iWEM.2016.7505030

5. Cerna RD, Yarleque MA (2018) A 3D compact wideband 16×16 butler matrix for 4G/3G applications. IEEE/MTT-S international microwave symposium—IMS 2018:16–19. https://doi.org/10.1109/MWSYM.2018.8439542

6. Tajik A, Shafiei Alavijeh A, Fakharzadeh M (2019) Asymmetrical 4×4 butler matrix and its application for single layer 8×8 butler matrix. In: IEEE transactions on antennas and propagation, vol 67, no 8, pp 5372–5379. https://doi.org/10.1109/TAP.2019.2916695

7. Adamidis G, Vardiambasis I (2006) Smart antenna design and implementation. A simple switched-beam antenna array based on a 8×8 butler-matrix network. In: 10th WSEAS international conference on communications (CSCC'06)

8. Yang Q, Ban Y, Zhou Q, Li M (2016) Butler matrix beamforming network based on substrate integrated technology for 5G mobile devices. In: 2016 IEEE 5th Asia-Pacific conference on antennas and propagation (APCAP), pp 413–414. https://doi.org/10.1109/APCAP.2016.7843268

9. Yang Q, Ban Y, Lian J, Zhong L, Wu Y (2017) Compact SIW 3×3 butler matrix for 5G mobile devices. Int Appl Comput Electromagnet Soc Symp (ACES) 2017:1–2

10. Abhishek A, Zeya Z, Suraj P, Badhai RK (2020) Design of beam steering antenna for 5G at 28GHz using butler matrix. In: 2020 5th international conference on computing, communication and security (ICCCS), pp 1–4. https://doi.org/10.1109/ICCCS49678.2020.9276492

11. Louati S, Talbi L, OuldElhassen M (2018) Design of 28 GHz switched beamforming antenna system based on 4×4 butler matrix for 5G applications. In: 2018 fifth international conference on internet of things: systems, management and security, pp 189–194. https://doi.org/10.1109/IoTSMS.2018.8554614

12. Yang Q, Ban Y, Yang S, Li M (2016) Omnidirectional slot arrays fed by stacked butler matrix for 5G handset devices. In: 2016 IEEE 9th UK-Europe-China workshop on millimeter waves and terahertz technologies (UCMMT), pp 245–247. https://doi.org/10.1109/UCMMT.2016.7874026

Pedestrian and Vehicle Detection for Visually Impaired People

Shripad Bhatlawande, Shaunak Dhande, Dhavanit Gupta, Jyoti Madake, and Swati Shilaskar

Abstract This paper proposes a method for helping visually impaired people by detecting and classifying incoming obstacles into pedestrians and vehicles on the road. When walking on roads or pathways, visually impaired persons have limited access to information about their surroundings; therefore, recognizing incoming pedestrians or cars is critical for their safety. Walking from one location to another is one of the most difficult tasks for visually impaired persons. White canes and trained dogs are the most often utilized instruments to assist visually impaired people in traveling and navigating. Despite their popularity, these technologies cannot offer the visually impaired people all of the information and functionality that persons with sight have access to and these aids are not that good for safe mobility. This proposed model aims to help visually impaired people by solving this problem. This paper proposes a model that can detect and classify pedestrians and vehicles on road using machine-learning and computer vision approaches. The system compares different classifiers based on SIFT and ORB feature extraction to evaluate the best approach. Different classifiers such as Random Forest, Decision Trees, SVM (3 kernels), and KNN are compared on the basis of testing accuracy, $F1$ score, recall, precision, sensitivity, and specificity. This study concluded that Random Forest yields the best result with 87.58% testing accuracy with SIFT feature extraction.

S. Bhatlawande (✉) · S. Dhande · D. Gupta · J. Madake · S. Shilaskar
Department of Electronics and Telecommunication, Vishwakarma Institute of Technology, Pune 411037, India
e-mail: shripad.bhatlawande@vit.edu

S. Dhande
e-mail: Shaunak.dhande19@vit.edu

D. Gupta
e-mail: dhavanit.gupta19@vit.edu

J. Madake
e-mail: jyoti.madake@vit.edu

S. Shilaskar
e-mail: swati.shilaskar@vit.edu

1 Introduction

In today's world, when we are experiencing a technological revolution, recent advances in the fields of computer vision and machine learning have provided us with tools to assist those who are in need. It is our responsibility to make use of these technologies to assist those people, and this project aims to assist visually impaired people because vision is the most important source of information for the typical person, and it is also the most basic requirement for anyone to live a regular life. Sight is one of the five primary senses (the others being hearing, touch, taste, and smell). You mostly obtain information through sight in your daily encounters, and those people who do not have proper vision or are completely blind face a lot of problems in their day-to-day life. Also factors like lack of information about their surroundings and problems faced in mobility are frequently leading blind people to isolation.

Many people have problems with their vision at some time in their life. Some people have lost their capacity to see distant objects. Others have trouble reading small text. The most common diagnosis for this group of individuals is visual impairment. Experts to characterize any sort of vision loss, whether total or partial blindness, use the phrase 'visual impairment.' Some people are completely blind, while others have a condition called legal blindness. They have not completely lost their vision, but it has deteriorated to the point that they must stand 20 feet away from an item to see it as clearly as someone with excellent vision could from 200 feet away.

According to the different surveys, there are 285 million visually impaired people worldwide, and among these, 39 million people are completely blind, and this number can be more in 2021. Also, large percentage of completely blind people are students. The number of blind persons over the age of 60 is rising at a rate of 2 million every decade. According to the Mospi, 2.8 million individuals suffer from vision diseases, according to comparison estimates based on the NSSO 2002 survey (58th round) [1]. According to the 2001 census, nearly 21 million individuals in India suffer from some form of impairment. According to the 2001 census, almost 21 million people in India have some kind of disability. This corresponds to 2.1% of the total population. Males account for 12.6 million of the country's handicapped, while females account for 9.3 million. 48.5% of people have vision problems. According to the JICA data profile on disability, 10.32% of Indians have vision problems [2].

Gender: Women account for 55% of the world's 253 million visually handicapped persons (139 million). A variety of variables contribute to difference in number of visually impaired people in each gender. Age is a main factor when it comes to sight disorders. According to many surveys, women life expectancy is better than men's life expectancy. This is the reason for more women with sight disorders as their number is more in old age group. Women in some nations are at a disadvantage

when it comes to accessing eye health care. Multiple socio-economic and cultural variables have contributed to this.

Socio-economic status: Low- and middle-income nations account for 89% of visually impaired persons. Asian countries add the most to the percent of visually impaired people in the world which is 62% of the total population of visually impaired people is in Asia only. A number of blind people are in n different parts of Asia, Southeast Asia has 24 million visually impaired people, East Asia has 59 million visually impaired people, and the number is highest in South Asia that is 73 million. On the other hand, countries with the highest income in the world have a very less percentage of visually impaired people.

Several solutions have been developed to assist visually impaired persons and improve their quality of life [3]. Unfortunately, the majority of these systems have limitations.

A variety of designs has been designed to fulfill the needs of visually impaired people and to enhance their living conditions. The approach followed by different researchers in the last three decades was to develop tools that used sensor-based technology for helping blind people with mobility and providing them safe while walking. Different tools developed to assist visually impaired people were the C-5 Laser Cane [4], the Nottingham Obstacle Detector [5], the Sonic guide [6], the people sensor [7], and Mowat Sensor [8], which are examples of these devices, which are collectively known as Electronic Travel Aids (ETAs). These ETAs, on the other hand, have not been widely adopted by their intended users, owing to the limited value of this category of systems [9]. Apart from technologies like audiobooks and other technologies in which sound is used to help the visually impaired, other technologies are also being researched. These technologies are specially designed to help visually impaired people. Wong et al. discussed project NAVI [10] which captures an image with help of a vision camera and processes it to stereo sound. There are few other papers with traditional approaches for helping the visually impaired like using sensors for detecting objects and the distance of an object in the path of a visually impaired person. For example, a guide cane sounds very similar to a normal cane but is different as it is developed to assist blind or visually impaired individuals in navigating barriers and other dangers securely and swiftly. The user pushes the button of the cane during operation. When the ultrasonic sensors on the cane identify an obstruction, it alerts the user. The cane and the motor are guided by an inbuilt computer that determines the best direction of travel [11]. Another effort is Blind People's Voice-Based Navigation System which uses ultrasonic sensors. The major goal of this project is to enable blind people to explore their surroundings independently [12]. To aid visually impaired people and improve their quality of life, several solutions have been created. The bulk of these systems, unfortunately, has flaws. Another study in this sector is an intelligent gadget that is designed for visually impaired persons to help them go to their destination securely and without incident. It is made up of a Raspberry Pi and PIC controller, a Global Positioning System (GPS), as well as sensors such as ultrasonic and other complementary sensors, and an Android-based application (APP) [13]. The Sound of Vision system [14] is a wearable sensor replacement device

that aids visually impaired persons in navigating by producing and transmitting an aural and tactile representation of their surroundings. Both audible and haptic feedback will be provided to the user. The system's usability and accuracy both need to be improved [15].

Other than the sensor-based approach, another approach is there to help visually impaired people which are the use of computer vision and artificial intelligence. Mekhalfi, Mohamed et al. discussed a study report on using smart technology to help blind persons recover their vision in interior surroundings, which aimed to make a prototype using which blind person can walk independently and (ii) recognize different items in public interior situations, which sought to satisfy the aforementioned demands. It combines a compact integrated gadget that may be worn on the chest with lightweight hardware components (camera, IMU, and laser sensors). Its methodologies are mostly based on advanced computer vision and machine-learning techniques. The user and the system interact via speech recognition and synthesis components [16]. Another way to detect obstacles is by using deep neural networks as discussed in [17, 18]. Another study offered a method that recovers the visual system's core role of object identification. The local feature extraction notion is used in this technique. The SFIT method and key points matching simulation results revealed strong object detection accuracy [19]. Another research proposes to design assistive technologies for visually impaired persons that will let them move more easily and increase their social inclusion. The majority of them aim to tackle navigation or obstacle avoidance problems, while some focus on assisting visually impaired persons in recognizing their surroundings. However, only a small percentage of individuals combine both skills [20]. Another study proposes an ambient intelligence system for blind people, which assists with a variety of scenarios and positions in which blind individuals may find themselves. The designed system identifies people approaching blind persons [12]. Another project aims to help blind people by making a smart gadget, which can distinguish faces, colors, and different things using artificial intelligence and picture processing. The detecting procedure is exhibited by the visually impaired individual being notified via sound or vibration [21]. Another study to help the visually impaired by detecting pedestrians is discussed in [22, 23], and for classification, SVM algorithm is used. But for helping visually impaired people navigate, detecting vehicles is also very important.

Another approach is a Smartphone-based approach like an application designed to assist visually impaired people inside, and it also provides feedback and detects landmarks nearby [24]. Lin et al. [25] discussed an application that can help visually impaired people in navigating but it also gives the option of offline and online modes. Another navigation system in phones for helping blind people with voice is discussed in [26], and navigation is one of the major issues faced by blind people [27]. When the system is turned on, the Smartphone takes a picture and transmits it to the server for processing. To distinguish different barriers, the server employs deep learning algorithms [28, 29]. The system's primary drawbacks are its high-energy consumption and the requirement for high-speed network access. The Tactile Way finder [29] is made of a tactile belt and a Personal Digital Assistant (PDA) that runs a Way finder app. The app keeps track of your position and path. The information is provided to

the tactile display when the travel direction has been determined. The vibrators in the belt can provide information to the user about navigation directions. Another study uses voicemail architecture to assist visually impaired persons in accessing e-mail and other multimedia aspects of the operating system (songs, text). SMS may also be read by the system in a mobile application [30]. Zhou et al. described the development of a low-cost assistive mobile application aimed at improving the quality of life of visually impaired persons. Other papers using Smartphone applications for helping blind people are [31, 32].

Other initiatives include a gadget that can supply blind people with distance information and make searching easier; in addition to identifying medications, the device can also provide the user with an audio signal to help them reach the desired object as quickly as possible. For this, RFID technology is used [33].

This project is solving problems of visually impaired people with a computer vision and machine-learning approach because the present sensor-based solutions' biggest flaw is its sensor performance. These sensors do not detect important elements such as sidewalk boundaries. The most promising technique to solve this challenge appears to be computer vision. Other applications of computer vision, such as localization, are possible. This project aims to help visually impaired people by accurately detecting and classifying incoming pedestrians and vehicles. This will solve one of the major problems faced by blind people which is mobility, and it will ensure their safety while traveling from one place to another. For the detection and classification of pedestrians, vehicles, and empty roads, we have created our own dataset. It consists of 12,920 images divided into 6460 positive images of 2 classes that are pedestrians and vehicles and 6460 negative images of empty roads. The resolution of images is 400×400. This paper covers the design and methodology of this project in Sect. 2, and the results of this project are in Sect. 3.

2 Methodology

This paper presents an object classification and detection system. The system detects (I) pedestrian and (II) vehicles. The block diagram of the system is shown in Fig. 1. The system consists of a camera, processor-based system, and earphones.

Fig. 1 Block diagram of system
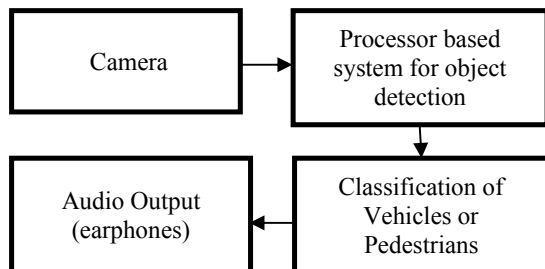
The camera captures the surrounding environment and provides it to the processor-based system. The processor-based system classifies and detects pedestrians or vehicles on the road. The system returns audio output via earphones to notify the visually impaired.

2.1 Data Acquisition and Pre-processing

The dataset includes 12,920 images captured from Indian roads. The images were captured from Samsung Galaxy M31 and Samsung Galaxy M40 with 32MP resolution each from various scenarios. The authors collected 70% of the dataset, and the remaining 30% was obtained from the Internet. The distribution of images in the dataset is presented in Table 1.

The images were resized to 280 × 430 pixels. The resized images were then converted to grayscale. Sample images from the dataset are shown in Fig. 2.

Table 1 Details of the dataset

Sr. No.	Class	Number of images
1	Vehicle	3230
2	Pedestrians	3230
3	Empty road	6460
	Total	12,920

Fig. 2 Sample images from the dataset

2.2 Feature Extraction

The model presents feature extraction from two descriptors. Scale-invariant feature transform (SIFT) and Oriented FAST and Rotated BRIEF (ORB) were used for feature extraction as shown in Fig. 3 and algorithm 1. A similar approach was used for the extraction of SIFT features.

ORB is a partial scale invariant descriptor. It detects key points at each level at a different scale. After finding the key points, the ORB assigns a shape to each key point such as left or right view depending on how the intensity levels change at that key point. The ORB uses the intensity centroid to find the change in intensity. The intensity centroid assumes that the stiffness of the angle is removed in the center, and this vector is used to identify the shape [34].

The moments of patch are defined as

$$m_{pq} = \int x^p y^q I(x, y) \tag{1}$$

In Eq. (1), (x, y) are the pixel coordinates and $I(x, y)$ is their gray value where (p, q) can take values of [0, 1].

The centroid of the image block is calculated using these moments as shown below

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \tag{2}$$
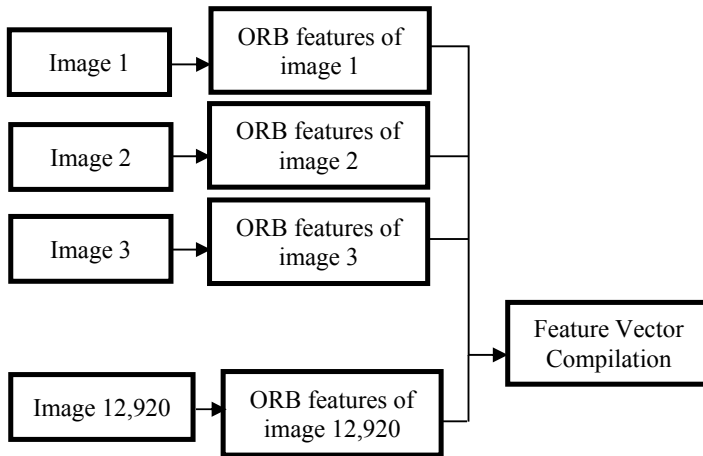
Fig. 3 Feature vector compilation using ORB

In Eq. (2), C is the centroid of image block and m_{10}, m_{01}, m_{11}, and m_{00} are the moments of respective pixels. The ORB is a computationally fast and efficient feature descriptor. It provided a feature vector of size 5,361,192 $\times$ 32.

Scale-invariant feature transform (SIFT) was the second feature descriptor. SIFT extracts features that ensures its identification even if the image orientation changes in the future. SIFT requires more time and memory as compared to ORB. The SIFT provided a feature vector of size of 5,625,904 $\times$ 128.

ORB and SIFT descriptors were used to compare and evaluate the best fit for the proposed system. All the extracted features were saved in a csv file individually for further processing.

Algorithm 1: Feature extraction using ORB

Input: Images (12,920)
Output: Feature vector (5,361,192 $\times$ 32)
  Initialization
1. For each class in data **do**
2. Specified path of dataset
3. For loop for each image in dataset
4. Image resize (280 $\times$ 430)
5. Conversion in grayscale
6. Feature extraction
7. Append in csv
8. For end
9.   Append csv's
10. For end
11. Final feature vector (5,361,192 $\times$ 32)

2.3 Dimensionality Reduction

SIFT and ORB provided large-size feature vectors. The dimensionality reduction approach was used for reducing the feature vector size. The initial ORB feature vector size was 5,361,192 $\times$ 32, and SIFT feature vector size was 5,625,904 $\times$ 128. The SIFT and ORB feature vectors were reduced using K-means clustering and principal component analysis (PCA). The ORB feature vector was divided into five clusters using K-means clustering ($k = 5$). The SIFT feature vector was divided in seven clusters using K-means clustering ($k = 7$). The number of clusters was decided based on the elbow plot. The acquired ORB and SIFT feature vectors were then transformed in five bin and seven bin histograms, respectively. The ORB and SIFT features for every image were then predicted using the pre-trained K-means model. The ORB feature vector size obtained was 12,245 $\times$ 5. The SIFT feature vector size obtained was 12,245 $\times$ 7. These input feature vectors were further optimized using principal component analysis (PCA). The number of components used was based on the maximum explained variance. This approach optimized the feature vector to size 12,245 $\times$ 3 and 12,245 $\times$ 5 for ORB and SIFT, respectively. The steps taken for dimensionality reduction for the ORB descriptor are shown in Fig. 4 and Algorithm 2. Similarly, the features extracted by SIFT were dimensionally reduced.

Fig. 4 Dimensionality reduction of feature vector
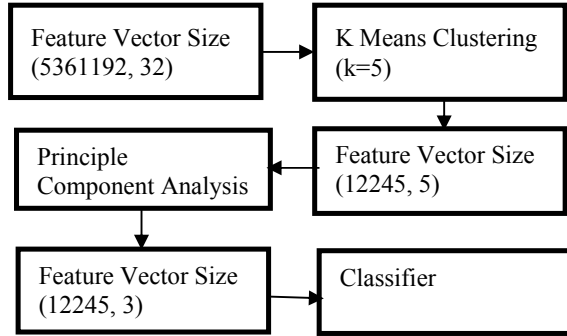
Algorithm 2: Algorithm for dimensionality reduction

Input: Feature vector (5,361,192 × 32)
Output: Reduced feature vector size (12,245 × 3)
    Initialization:
1. K-Means clustering on ORB feature vector ($k = 5$)
LOOP Process
2. for every image in specified path **do**
3. Perform k-means prediction
4. Applying normalization
5. Append in csv to form feature vector
6. end for
7. Standardization using standard scalar
8. Performing PCA (n components $= 3$)
9. PCA transform
10. Reduced feature vector size
11. return optimized feature vector (122,453)

2.4 Classification of Vehicles and Pedestrian

The resultant feature vectors were given as an input to an array of classifier algorithms. The system uses four different classifiers to evaluate the best fit for the model, namely: (i) Random Forest, (ii) Support Vector Machine, (iii) K-Nearest Neighbors, and (iv) Decision Tree as shown in Fig 5.

The decision tree was the first classifier. The decision tree algorithm is closest to how the human brain works. In the decision trees, different attribute selection measures are used to decide the position of the node.

The second classifier used was SVM. This classifier creates a line that separates the data into classes. SVM uses set of functions that are defined as kernels. The system uses four SVM kernels, namely: (i) Linear, (ii) RBF, (iii) Polynomial, and (iv) Sigmoid. The polynomial hyperplane equation is
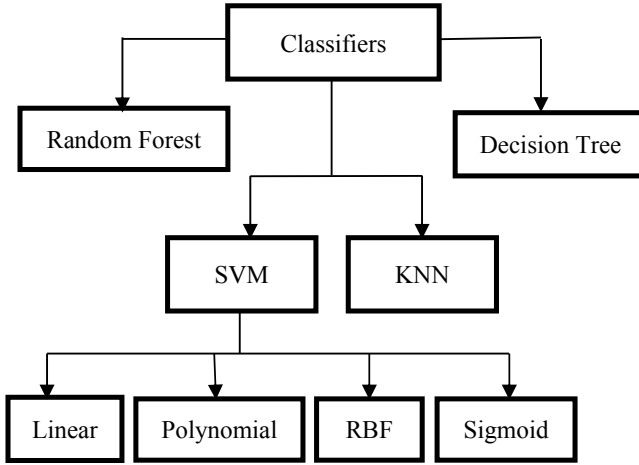
Fig. 5 Different classifiers used to predict class in the proposed system

$$k(x_i, x_j) = (x_i * x_j + 1)^d \tag{3}$$

'd' is the degree of a polynomial, 'x_i' is the input, and 'x_j' is the support vector in Eq. (3).

The third classifier used was KNN. KNN algorithm classifies the data with respect to its neighboring data classes. The neighbors are to be considered depending on the value of k ($k = 5$) in this case. The distance 'd' in KNN is calculated using the Euclidean distance function. The distance is calculated as below:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{4}$$

(x_1, y_1) and (x_2, y_2) are the coordinates of the two points in Eq. (4).

Random Forest was the fourth and last classifier used in this system. Random Forests are addition of multiple decision trees. It estimates the maximum vote to predict the classification. To know the impurity of a particular node, Random Forests have a concept called Gini Index. The Gini Index is calculated as follows where P + and P − are the probabilities of positive and negative class.

$$\text{G.I.} = 1 - \left[(P+)^2 + (P-)^2\right] \tag{5}$$

The prediction by a classifier is sent to the user in the form of audio. Using the 'Pygame' library in python, an audio message with respect to the output will notify the visually impaired. The complete flow of system methodology is shown in Fig. 6.
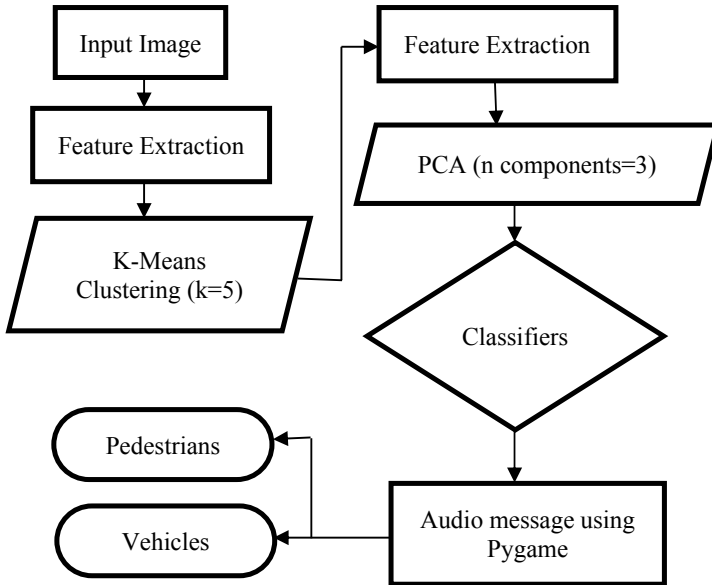
Fig. 6 Overall flow of proposed system

3 Results and Discussion

The dataset was split into 80% training data and 20% testing data. Four supervised classification algorithms were used for classification. The SIFT and ORB models were compared based on the testing accuracies of these classifiers. Random Forest had the greatest testing accuracy of 82.03% for the ORB-based system. The decision tree had a testing accuracy of 75.21%. The testing accuracy of SVM Linear was 77.05%, SVM Sigmoid was 56.18%, SVM RBF was 77.05%, and SVM Polynomial was 71.82%. The testing accuracy of KNN was 75.74% for '$k = 5$.' Random Forest had the highest accuracy of 87.58% for the SIFT-based system. The decision tree had a testing accuracy of 82.64%. The testing accuracy of SVM Linear was 82.85%, SVM Sigmoid was 65.41%, SVM RBF was 85.05%, and SVM Polynomial was 80.97%. The testing accuracy of KNN was 82.52% for '$k = 5$.' The accuracy scores for ORB and SIFT-based models are compared in Table 2. The comparison of these classifiers based on performance parameters such as $F1$ score, recall score, and precision is shown in Table 3 for ORB and SIFT-based models.

After evaluating the accuracies of all four classifiers based on performance parameters, Random Forest had the highest accuracy for features extracted by SIFT descriptor. Though its accuracy score is nearly 88%, it can be increased by improving the quality and quantity of the dataset.

Table 2 Training and testing accuracies for ORB and SIFT

Classifier	Training accuracy		Testing accuracy	
	ORB (%)	SIFT (%)	ORB (%)	SIFT (%)
Random forest	96.45	98.34	82.03	87.58
Decision tree	92.35	94.37	75.21	82.64
SVM sigmoid	57.28	64.06	56.18	65.41
SVM RBF	77.97	85.11	77.05	85.05
SVM Poly	71.77	80.48	71.82	80.97
SVM linear	77.43	82.14	77.05	82.85
KNN	82.23	88.54	75.74	82.52

Table 3 Performance Parameters for ORB and SIFT

	ORB descriptor			SIFT descriptor		
Classifier	F1 score (%)	Recall (%)	Precision (%)	F1 score (%)	Recall (%)	Precision (%)
Random forest	80.00	79.98	80.03	87.02	86.96	87.21
Decision tree	73.08	73.23	73.08	82.25	82.67	82.22
SVM sigmoid	52.26	52.73	52.97	65.27	65.55	65.27
SVM RBF	75.19	75.90	74.76	84.64	84.74	84.91
SVM poly	68.36	66.86	71.54	79.45	77.41	82.96
SVM linear	75.09	75.80	74.62	82.38	82.65	82.36
KNN	73.80	74.27	73.50	82.07	82.36	82.08

4 Conclusion

The proposed system gives the best accuracy for a Random Forest classifier with SIFT descriptor. The testing accuracy for the Random Forest is 87.58%, which is a very good accuracy. After comparing all the performance parameters, we can conclude that SIFT is the best descriptor for this proposed system. Thus, it is fit to assume that the proposed model can be used for further applications as it has good accuracy. The proposed model can classify the given object into pedestrians or vehicles. This ability would be extremely useful for helping blind people. When integrated with external hardware, the proposed system can be used for alerting the blind person about incoming pedestrians and vehicles. Blind people can be notified by vibration or sound when in front of incoming objects. There can be unique outputs for pedestrian and vehicle objects so that the person knows the object type. The proposed system has few limitations too which can be improved in future. The detection and classification

of objects could be incorrect in a few instances. The distance at which the object is recognized may vary. All these limitations can be improved when the proposed system is modified and equipped with appropriate hardware. The proposed model can be utilized in the domain of autonomous vehicles as well. Autonomous vehicles can use this system to detect incoming or stationary objects and avoid them. This model is trained on images of Indian origin. Thus, compared to other models, the proposed system will have more edge with respect to accuracy and detection. This system ensures that visually impaired people travel and navigate independently without taking someone's help, and it will ensure their safety while walking on pathways or roads. The collection of Indian origin datasets is also a major contribution displayed in this paper.

References

1. Kumar S, Sanaman G (2013) Preference and use of electronic information and resources by blind/visually impaired in NCR libraries in India. J Inf Sci Theory Pract 1(2):69–83
2. Elmannai W, Elleithy K (2017) Sensor-based assistive devices for visually-impaired people: current status, challenges, and future directions. Sensors 17(3):565
3. Kaczmirek L, Wolff KG (2007) Survey design for visually impaired and blind people. In: International conference on universal access in human-computer interaction. Springer, Berlin, pp 374–381
4. Benjamin JM, Ali NA, Schepis AF (1973) A laser cane for the blind. In: Proceedings of the San Diego biomedical symposium, vol 12, no 53–57
5. Bissitt D, Heyes AD (1980) An application of bio-feedback in the rehabilitation of the blind. Appl Ergon 11(1):31–33
6. Kay L (1974) A sonar aid to enhance spatial perception of the blind: engineering design and evaluation. Radio Electron Eng 44(11):605–627
7. Ram S, Sharf J (1998) The people sensor: a mobility aid for the visually impaired. In: Digest of papers. Second international symposium on wearable computers (Cat. No. 98EX215). IEEE, pp 166–167
8. Velazquez R, Pissaloux EE, Guinot JC, Maingreaud F (2006) Walking using touch: design and preliminary prototype of a non-invasive ETA for the visually impaired. In: 2005 IEEE engineering in medicine and biology 27th annual conference. IEEE, pp 6821–6824
9. Blasch BB, Long RG, Griffin-Shirley N (1989) National evaluation of electronic travel aids for blind and visually impaired individuals: Implications for design. In: RESNA 12th annual conference, pp 133–134
10. Wong F, Nagarajan R, Yaacob S (2003) Application of stereovision in a navigation aid for blind people. In: Fourth international conference on information, communications and signal processing, 2003 and the fourth Pacific Rim conference on multimedia. Proceedings of the 2003 joint, vol 2. IEEE, pp 734–737
11. Ulrich I, Borenstein J (2001) The GuideCane-applying mobile robot technologies to assist the visually impaired. IEEE Trans Syst Man Cybern-Part A Syst Hum 31(2):131–136
12. Hudec M, Smutny Z (2017) RUDO: a home ambient intelligence system for blind people. Sensors 17(8):1926
13. Chang Y-H, Sahoo N, Lin H-W (2018) An intelligent walking stick for the visually challenged people. In: 2018 IEEE international conference on applied system invention (ICASI). IEEE, pp 113–116
14. Caraiman S, Morar A, Owczarek M, Burlacu A, Rzeszotarski D, Botezatu N, Herghelegiu P, Moldoveanu F, Strumillo P, Moldoveanu A (2017) Computer vision for the visually impaired:

the sound of vision system. In: Proceedings of the IEEE international conference on computer vision workshops, pp 1480–1489

15. Huang C-Y, Wu C-K, Liu P-Y (2022) Assistive technology in smart cities: a case of street crossing for the visually-impaired. Technol Soc 68:101805

16. Mekhalfi ML, Melgani F, Zeggada A, De Natale FGB, Salem MAM, Khamis A (2016) Recovering the sight to blind people in indoor environments with smart technologies. Expert Syst Appl 46:129–138

17. Kumar A, Sai Satyanarayana Reddy SS, Kulkarni V (2019) An object detection technique for blind people in real-time using deep neural networks. In: 2019 fifth international conference on image information processing (ICIIP). IEEE, pp 292–297

18. Castillo-Cara M, Huaranga-Junco E, Mondragón-Ruiz G, Salazar A, Barbosa AO, Antúnez EA (2016) Ray: smart indoor/outdoor routes for the blind using Bluetooth 4.0 BLE. Proc Comput Sci 83:690–694

19. Jabnoun H, Benzarti F, Amiri H (2015) Object detection and identification for blind people in video scene. In: 2015 15th international conference on intelligent systems design and applications (ISDA). IEEE, pp 363–367

20. Bai J, Liu Z, Lin Y, Li Y, Lian S, Liu D (2019) Wearable travel aid for environment perception and navigation of visually impaired people. Electronics 8(6):697

21. Al-Muqbali F, Al-Tourshi N, Al-Kiyumi K, Hajmohideen F (2020) Smart technologies for visually impaired: assisting and conquering infirmity of blind people using AI technologies. In: 2020 12th annual undergraduate research conference on applied computing (URC). IEEE, pp 1–4

22. Kang S, Byun H, Lee S-W (2003) Real-time pedestrian detection using support vector machines. Int J Pattern Recognit Artif Intell 17(03):405–416

23. Kang S, Byun H, Lee S-W (2002) Real-time pedestrian detection using support vector machines. In: International workshop on support vector machines. Springer, Berlin, pp 268–277

24. Sato D, Oh U, Naito K, Takagi H, Kitani K, Asakawa C (2017) Navcog3: an evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In: Proceedings of the 19th international ACM SIGACCESS conference on computers and accessibility, pp 270–279

25. Lin B-S, Lee C-C, Chiang P-Y (2017) Simple smartphone-based guiding system for visually impaired people. Sensors 17(6):1371

26. Lima A, Mendes D, Paiva S (2018) Outdoor navigation systems to promote urban mobility to aid visually impaired people. J Inf Syst Eng Manag 3(2):14

27. Alwi SR, Wan A, Noh Ahmad M (2013) Survey on outdoor navigation system needs for blind people. In: 2013 IEEE student conference on research and development. IEEE, pp 144–148

28. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788

29. Manduchi R, Coughlan J (2012) (Computer) vision without sight. Commun ACM 55(1):96–104

30. Nilesh J, Alai P, Swapnil C, Bendre MR (2014) Voice based system in desktop and mobile devices for blind people. Int J Emerg Technol Adv Eng (IJETAE) 4(2):404–407

31. Zhou D, Yang Y, Yan H (2016) A smart "virtual eye" mobile system for the visually impaired. IEEE Potentials 35(6):13–20

32. Paneels SA, Varenne D, Blum JR, Cooperstock JR (2013) The walking straight mobile application: helping the visually impaired avoid veering. Georgia Institute of Technology

33. Dionisi A, Sardini E, Serpelloni M (2012) Wearable object detection system for the blind. In: 2012 IEEE international instrumentation and measurement technology conference proceedings. IEEE, pp 1255–1258

34. Tyagi D (2019) Introduction to ORB (oriented FAST and rotated BRIEF)
35. Mataró TV, Masulli F, Rovetta S, Cabri A, Traverso C, Capris E, Torretta S (2017) An assistive mobile system supporting blind and visual impaired people when are outdoor. In: 2017 IEEE 3rd international forum on research and technologies for society and industry (RTSI). IEEE, pp 1–6

Measuring Effectiveness of CSR Activities to Reinforce Brand Equity by Using Graph-Based Analytics

Krishna Kumar Singh and **Aparajita Dasgupta Amist**

Abstract Although CSR activities are not new for the industrialists, government of India implemented stick rule for industries to take part in social causes under corporate social responsibility (CSR). After decades of implementation, industries are showing positive responses towards CSR activities to contribute for the society. Investments are growing day by day. There is dearth of model in the annals which can find out real effectiveness of their investment on the society at large. Industries are curious to know how their contribution helping for the company's business. In this paper, researchers tried to build a model to understand effectiveness of CSR in terms of qualitative as well as quantitative parameters. Under qualitative measures, authors tried to measure how CSR contributed in enhancement of brand value as well as positivity towards products of the company. Quantitatively authors tried to measure how many times or at what rates general people change their mindset to purchase products of a brand by getting influenced from CSR news on public platforms. Authors used mining and analytical tools like anaconda orange and Neo4J to build model and calculate impact. A quantitative method has been developed on the basis of whole procedure as well as formulae for calculating quantitative impact of CSR activities of a company on the basis of information shared by company in public domain. As a result, researchers found that there is short-term as well as long-term positive correlation between companies' growth and CSR information shared in the public domain. Real impact of these activities depends from sector to sector and company to company depending upon various factors as discussed in the paper. Due to the lack of data on account of privacy laws of the land, research has been done in control manner with limited data set.

Keywords Data mining · Graph-based analytics · Data mining · Corporate social responsibility (CSR) · Neo4J

K. K. Singh (✉)
Symbiosis Centre for Information Technology, Pune, India
e-mail: krishnakumar@scit.edu

A. D. Amist
Amity Global Business School, Amity University, Noida, India
e-mail: adgupta@amity.edu

1 Introduction

Corporate social responsibility or CSR is not necessarily by giving a one-time amount as a checkbook charity as it is called, but to engage with the community, engaged with the best minds in one's company. Lua et al. [1] indicated that nowadays the government, the bureaucracy and businesses can collaborate in a holistic way to ensure this is done because these are two sides of the same coin with the society as the primary beneficiary of every activity they undertake. Corporate social responsibility or CSR in India is a mandatory law. CSR adopted by trusted branded companies in India Fernandes et al. [2] such as Infosys Ltd., Mahindra and Mahindra Ltd., Tata Chemicals Ltd., ITC Ltd., Vedanta Ltd., Wipro Ltd., Hindustan Unilever Ltd., Godrej Consumer Products Ltd., Grasim Industries Ltd. and Hindustan Petroleum Corporation Ltd. is built around the strategy of how an organisation treats its people, the environment in which they operate in and the communities they serve and all of the different facets that create a sustainable business. In a lot of ways, CSR is becoming expected of businesses. Nowadays, employees and both consumers are demanding it. The mindset of today's organisations is about understanding how employees interact with the company in a responsible and sustainable way, then it opens up the way for opportunities in the area on how to increase productivity, increase employee retention, and it doesn't have to be a cost centre Singh MK and Singh KK [3].

Today organisations when embarking on a CSR activity and strategy look into metric and measurable goals with definite timelines and goals Jiminez et al. [4]. When CSR is incorporated well into an organisation, it is an extension of the brand. Companies with a good CSR strategy align it with the core capabilities of their business and their people. A multinational corporation—a fast-moving consumer good company—an FMCG company which is into food brands and retailing making sure it aligns with what is right in terms of value offerings and its core business offerings given to the customers towards creating business partnerships. Due to the increased engagement activities undertaken by companies in CSR, the consumers attitude and perception of the companies' brand is positively reinforced through its product offerings [5, 6].

New launches of CSR campaigns have influenced consumer buying behaviour in a more direct and ethical way of conducting business, and there is no increase of cost and quality variation in the products and services. The negative impact and information of CSR campaigns affect all consumers, only those truly motivated towards social behaviour are affected by corporate initiatives related to the welfare of society [6]. Therefore, further investigation is needed to shed light on the differential responses consumers' shows when exposed to such initiatives resulting in the company's brand and reputation [7, 8]. Companies are working towards sustainability and economic development of society are conducting useful dialogues towards integration of activities resulting in meeting the demands and expectations, when reporting back their actions to generate real value for their different stakeholders. Hence, it is necessary to evaluate the impact of the actions undertaken by the companies and interpretations and feelings being generated towards every interest group, and so be ready to

redefine priorities and responsibilities within the organisation from a triple perspective: economic, social and environmental [9]. Social responsibility through CSR is a policy and practice adopted by companies who integrate social and environmental issues into the activities which are harmoniously integrated with the interests of the stakeholders guided by respect for the individual and society and the environment [9].

The social information of a company provides information on varied aspects of community-related towards social, community, ethical and environmental issues. Any kind of changes towards disclosure of information has been shared by some scientists that they relook at the given missing information and determine the factors of social reporting within the purview of the management's decision. While another aspect of research investigates the quality of the accounting statements and the relevant applied accounting principles. However due to the complexity, sensitivity and relativity towards the wide area of subject, there is some chance of finding it difficult to assess. It may be noted that all the information may not be applicable towards the quality of social information. The aim of the research is to understand how data science tools are used effectively to analyse the social content of data of socially responsible companies in India. Leading mobile operator companies, oil and petroleum companies, power companies, consulting companies and software IT-related companies in India are using Neo4j which help them to manage networks, control access and convert a customer 360. These graphic databases are used in Fraud detection, Real-time recommended engines, master data management (MDM), network and IT operations and identity and access management (IAD).

2 Literature Review

During the early 1970s and throughout the whole period of 1990s and 2000s, extensive study in operations techniques and analysis techniques started out to present to multivariate methods [10] such as correlational, multiple regression and discriminant analysis techniques. Later these were expanded to conjoint analysis and cluster analysis, multidimensional scaling (MDS) and canonical correlation analysis. Paul Green of Wharton College and his former college student significantly Vithala Rao proclaimed the use of joint analysis and metric multidimensional scaling. A new perspective was introduced into research and the discipline of marketing and research progressed from economics to psychometrics study. Recently, there has been a noticeable shift towards studies on large statistics (non-numerical statistics such as textual content messaging) on social platforms the usage of natural language processing (NLP) and artificial intelligence (AI). A unique recognition is the power of the word of mouth (WoM) creating an effect on all customers and influencers. In the digital era, social media has gained importance as a digital advertising medium with viewer/user/consumer-generated various types of content material. An example, of YouTube, it provides—hundred of thousands of users uploaded films every day. Reels

and short-video have gained primacy on the social platforms like WhatsApp, Facebook and Instagram. Research on this content through qualitative analysis/statistics has further modified the vocabulary from marketing models to marketing analysis and analytics [11].

Sensor technologies enable to capture volumes of huge data, that are beyond the capacity of traditional tools to capture and store. They are capable to collect and store real-time data like that of consumers' consumption, weather conditions and other various other criterion. Usage of big data with the advanced sensors technology has resulted into much-accurate decision-making. With big data, organisations can handle huge volume of data—structured as well as unstructured—which enables a better understanding and automation of processes [12, 13].

Big data can enhance the organisational performance through leading to more precise and effective data-based decisions and optimising various business processes. Advanced analytical tools enable better prediction of future activity and performance by utilising this data, e.g. load forecasting (LF). Thus, enables businesses to fine-tune processes for improved performances. A number of digital technologies such as SQL (NoSQL) have been developed to address the technical aspects of handling and processing large amounts of real-time data simultaneously and captured from various sources. In this paper, authors have utilised graphics technology as a big data tool, in order to store and analyse these volumes of data. Here graph represents a series of vertices and edges, in other words, a set of nodes indicating their relationships via connecting the nodes. In the graphs, nodes represent various entities and the connections represent the ways in which the entities relate to the world as relationships. Nodes and their relationships, both, are support properties, constituting the key-value pairs. In a graph database, data are stored in graph, and the records are called nodes. Relationships are represented through connecting the nodes [14].

In order to organise nodes into groups, a label is given to each group of related nodes. For this, one of the most accepted graph databases is Neo4j. Critical issue in a graph database is the way query is formulated. Like any other graph databases, Neo4j uses a powerful mathematical concept of graph traversal from graph theory and makes it as a robust and efficient engine for handling queries. Neo4j is fully compliant ACID database and based on query language—Cypher (CQL) as Robinson et al. [14]. Neo4j graph database is written in Java, which can handle around 34 billion of nodes and their relationships. Thus, it consists of nodes, properties, relationships, labels and data browser as its building blocks [15, 16].

3 Pre-processing of Data

Authors took data from nations CSR portal (National CSR Portal), government of India and social media account like Twitter and Facebook. As data is heterogeneous in nature and fetched from different sources, pre-processing of data is required. Data from national CSR portal is available in two-dimensional format in rows and columns, so only removal of errors and some more discrepancies removal will serve our purpose

but data from social media needs more attention because of its heterogeneity. For the purpose of current research, authors considered data related to identity and relation between identities (nodes) to know their patterns. API of these social media account is being used through python language to fetch data. Pre-processing of Twitter data is done with the help of anaconda orange (Fig. 2). Authors used graph to establish relationship and know about effectiveness of CSR activities with the help of Neo4J. Figure 1 shows creation of nodes for the graph containing all properties of a person.

Fig. 1 Creation of nodes
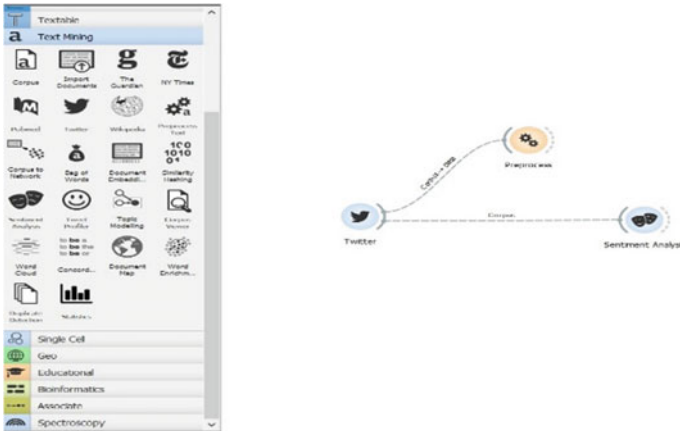
Fig. 2 Pre-processing of Twitter data

4   Modelling and Analysis

Authors took multiple data science approaches to analysis data available namely data visualisation, graph and algorithmic methods. Figure 3 shows growth chart of NIFTY in 27 years of time span. Nifty witness a journey from 2000 in 1996 to 18,000 in 2021 and rate of growth of Indian stock market in the last 25 years are more than 10%. After government of India regulation on compulsorily corporate social responsibility (CSR) for the industry, many of these companies have been coming forward to contribute voluntarily for this novel cause which helps society as a whole in different walks of life. Investments under corporate social responsibility (CSR) have been increasing year on year with the growth rate of 8–10%. A comparative study of rate of growth of nifty and investment in CSR in India shows impressive results and both are showing upward positive trends with 6–10% in long run (e.g. more than 25 years). To understand deeper into CSR scenario, authors tried to look into the leaders in the Indian industries like Reliance, Tata and Mahindra and Mahindra as case study. Data with many sources show that big players have greater pie in the CSR activities. Next figures show the growth of NIFTY in comparison with these big players. Figure 4 shows growth chart of NIFTY in respect to Reliance Ltd. in 26 years of time span. Figure 5 shows growth chart of NIFTY in respect to TCS Ltd. in 16 years of time span. Figure 6 shows growth chart of NIFTY in respect to Tata Chemical Ltd. in 16 years of time span. Figure 7 shows spending of Mahindra and Mahindra in 5 years of time span.
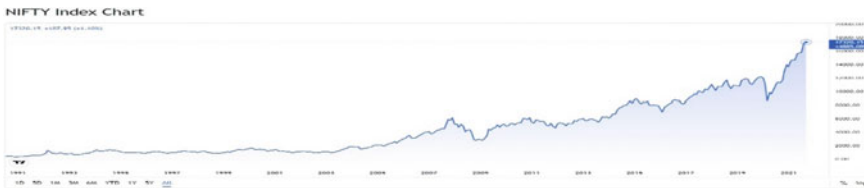
Fig. 3   Growth chart of NIFTY in 27 years
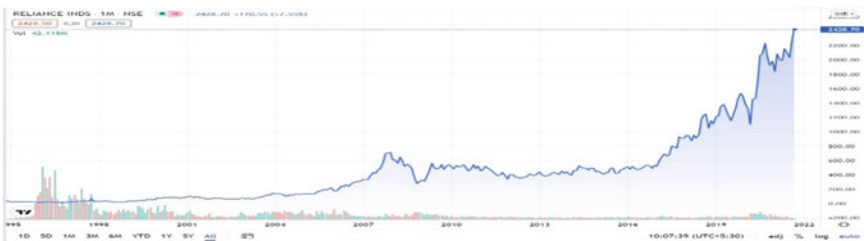
Fig. 4   Growth chart of NIFTY w.r.t Reliance Ltd. in 26 years
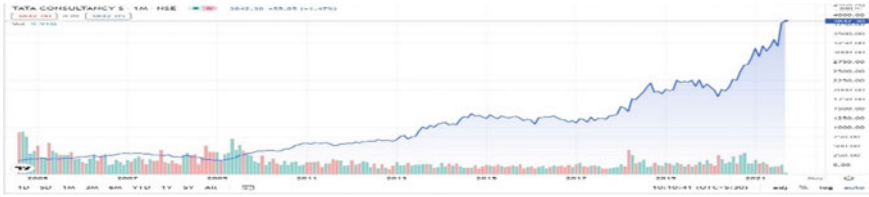
Fig. 5 Growth chart of NIFTY w.r.t TCS Ltd. in 16 years
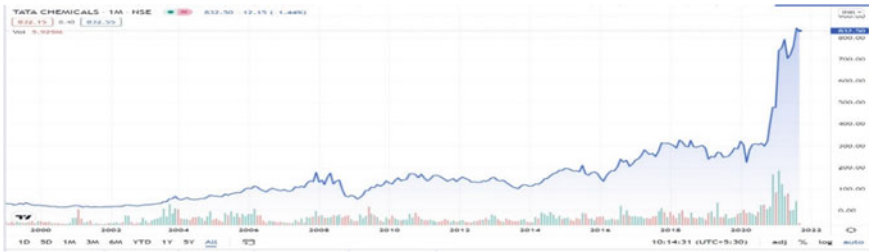
Fig. 6 Growth chart of NIFTY w.r.t Tata Chemical Ltd. in 16 years
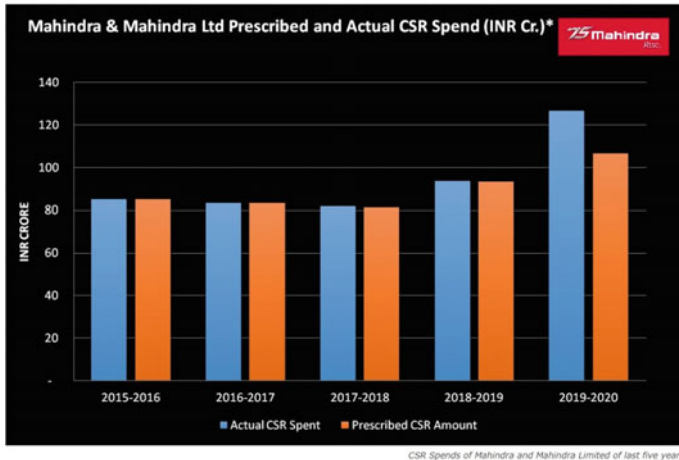
Fig. 7 Rate of CSR spending of M and M in 5 years. *Source* NIFTY index charts and quotes—trading view—India

By seeing comparative rate of growth of CSR spending by reliance, m&m, TCS and Tata steel with nifty authors can conclude that both have positive correlation. To understand impact of CSR information in the public domain, authors took data of social media account and tried to understand rate of growth of people reaction

on the post by using graph-based analysis (Neo4J). Figure 8 shows the graph of
connectivity of members (general public). S_1, S_2, S_3, ..., S_n are sources on which
post available, and P is public connected to the person. Model of the information
source and connectivity is shown in graph below. To implement the model displayed
in the graph, authors used social media data about information disseminated by
industry players and people reaction on the post. Authors also considered reactions
on reaction, which gives compounding multiplier impact of post on social media. In
this graph, people (p) have profiles and connected with other people or sources of
CSR information send. People connected to these sources are seeing, like/dislike or
share through its profile. Data-related each activity can be analysed by text analytical
methods. Once any person resends the same content, it becomes sub-sources. All
activities of previous sources, current sources and others are being added to see its
impact on the public.

Figure 9 shows nodes and relationship in graph using Neo4J. To trace these activ-
ities, graph method is most suitable. Each node of the graph becomes source (or
sender) of information (denoted as s) and people who see these become receiver of
information. Source and people who received information are nodes in the graph. All
information related to that person will save in nodes. When a person sends an infor-
mation to their friend or general public and other person reacted on the information,
then these relationships are shown as edge of the graph. Because of send and resend
of information, there are multiple relationships established among nodes and it can
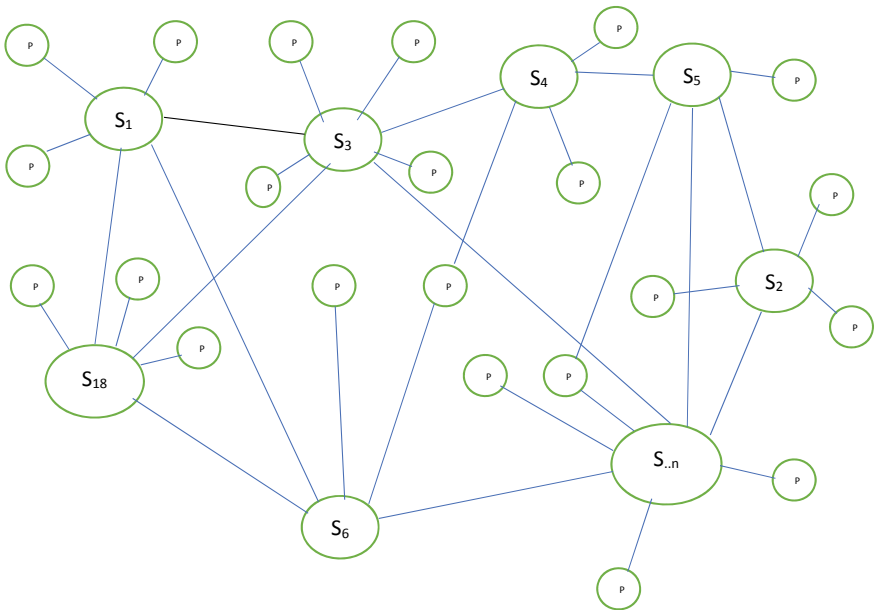
Fig. 8 Graph of connectivity on public platforms
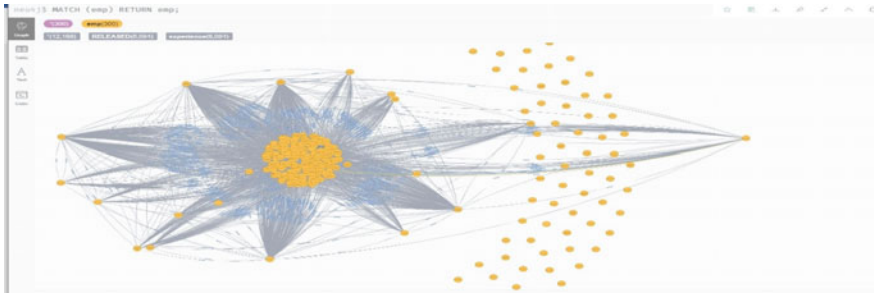
Fig. 9 Nodes and relationship in graph using Neo4J

be represented as edges of the graph. Apart from Neo4j, authors also used anaconda orange for pre-processing of data as well text analytics and social media analytics.

$S_1, S_2, S_3, \ldots, S_n$ are sources on which post available.

P is public connected to the person.

Figure 10 shows some of the tweets on CSR through companies' official Twitter account and individual account. Number of re-twits will show in the node of graph, and its impact will be calculated by mapping users who like or re-twit and purchase company products in limited time period. Graph method shows relationship among user's spending on product and their behaviour on CSR activities of respective company. Author's used anaconda orange to know positive responses and neo4j to map both activities simultaneously.
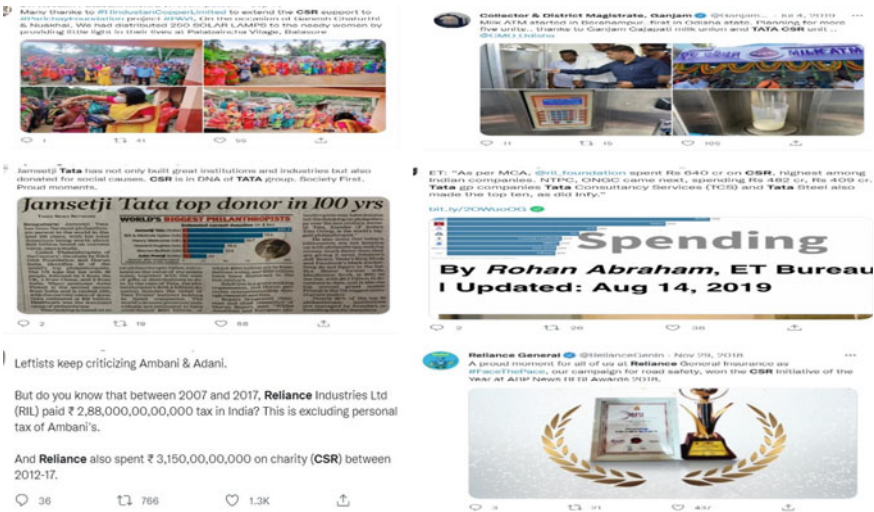
Fig. 10 List of various tweets on CSR and its impact as like, retweet or forward. *Source* Twitter

To understand and measure effectiveness of an information (media post) in public domain, below formulae has been established. These calculations are based on multiple factors like number of posts, number of likes on the post, direct positive feedback on the post, number of people in the connection list of the person liked and how many of them are reacting positively or negatively on the post, etc.

$$E_1^p = \left\{ L_n * F_n + \sum_n^1 S_1^{L+F} \right\}$$

E is the effectiveness from 1 to n post, where p is the positive feedback
L is the number of likes from n post
F is the number of family members or close connected people of account holder
S is the number of shares of post with exponent of number of likes and number of family members.

According to the above formulae, effectiveness (E) of a post is calculated by multiplication of likes and feedback on the post directly or indirectly. Plus, summation of number of times these posts has been shared. Each shared post has gain likes, feedbacks, etc., from direct or indirect channels. So, total effectiveness of the post will be calculated by adding all its short- or long-term impact. Multiplier effect of graph will graph relationship and will show spreading of information, and its impact will be calculated but the above formulae. There is multiple benefits of this model but we can consider two most important. One is enhancing brand value and other direct monetary benefit for the companies. Brand value enhancement will be seen by calculating number of times these posts are resend, people liked it or given positive feedback on public platform. But to calculate real impact in total conversion value of CSR, we have to integrate with the data of people and understand their purchasing patterns of the goods or services. If a person is influenced by CSR activity done by company, he/she will consider these companies' products as and when they required these products in near future. Because of regulatory limitations, authors have access of limited personal data. Implementation of model on the data available shows an impressive result of more than 8% conversion rate and 46% increase in the brand perception. With the help of the model, corporate houses pin pointedly parget its customer and able to know their grievances and expectations which further leads of increase faith in the products. Companies can also rebuild their image among general people for the long-term growth and development.

5   Results and Conclusion

Current model will measure effectiveness of CSR activity-related information in public domain and its impact on the performance of the companies. Relationship established among people's account and their reactions on information available

is a great source of data to measure effectiveness through formulae and algorithm suggested in the model. Shared data have computing and multiplier effect on the company's performance. Results on the tested data showed positive correlation between CSR activities and performance of the companies. Although because of limited data available with the authors, validation has been done in the closed format. If data will increase, then results will be more realistic to apply. Data related to user purchased patterns, discussion about products, dynamic shift from existing to new products in the segment after CSR information will provide more exact results. Integration of big data analytical tools on the data related to customer purchase on online sites and other movement will provide more realistic results. It helps in changing perception of the brand by re-enforcing brand equity using the above model.

References

1. Lua W, Wangb W, Leec H (2013) The relationship between corporate social responsibility and corporate performance: evidence from the US semiconductor industry. Int J Prod Res 51(19):5683–5695
2. Fernandes K, Thacker H (2020) Top 100 companies in India for CSR in 2020. https://thecsr journal.in/top-indian-companies-for-csr-in-2020/
3. Singh MK, Singh KK (2021) A review of publicly available automatic brain segmentation methodologies, machine learning models, recent advancements, and their comparison. Ann Neurosci 28(1–2):82–93
4. Jiminez JVG, Ruiz-de-Maya S, Lopez I (2017) The impact of congruence between the CSR activity and the company's core business on consumer response to CSR. Span J Market ESIC
5. Bhattacharya CD, Korschun D, Sen S (2009) Strengthening stakeholder-company relationships through mutually beneficial Corporate Social Responsibility initiatives. J Bus Ethics 85(2):257–272
6. Bhattacharya CB, Sen S (2004) Doing better at doing good, whey, why, and how consumers respond to corporate social initiatives. Calif Manage Rev 47(1):9–24
7. Vilaplana J, Vintro C, Pons A (2021) Impact of corporate social responsibility in mining industries. https://www.sciencedirect.com/science/article/pii/S0301420721001318?via%3Dihub
8. Vaaland T, Heide M, Gronhaug K (2008) Corporate social responsibility: investigating theory and research in the marketing context. Eur J Mark 42(9/10):927–953
9. Sapkauskiene A, Leitonienea S (2015) Quality of corporate social responsibility information. In: 20th international scientific conference economics and management-2015 (ICEM-2015)
10. Sheth J (2017) Genes, climate, and consumption culture: connecting the dots. Emerald Group Publishing
11. Sheth KCH (2021) New areas of research in marketing strategy, consumer behavior, and marketing analytics: the future is bright
12. Singh KK, Pandey R (2021) Predicting employee selection using machine learning techniques. In: 2021 3rd international conference on advances in computing, communication control and networking (ICAC3N), pp 194–198. https://doi.org/10.1109/ICAC3N53548.2021.9725777
13. Thomas Erl, Khattak W, Buhler P (2016) Big data fundamentals concepts, drivers & techniques, pp 19–42. ISBN-13: 978-0-13-429107-9
14. Robinson I, Webber J, Eifrem E (2015) Graph databases new opportunities for connected data; 2015, pp 171–211; Vukotic A, Watt N, Neo4j in action; 2015. ISBN: 9781617290763, pp 1–79

15. Arbër P, Daniela M, Lyudmila S (2017) Modeling and processing big data of power transmission grid substation using Neo4j. Proc Comput Sci 113:9–16
16. Chaudhary S, Singh V, Bhagat A (2020) The changing landscape of CSR in India during COVID-1. Strategic investment research unit (SIRU). The changing landscape of CSR in India during COVID-19

Design and Implementation of Built-In Self-Test (BIST) Master Slave Communication Using I2C Protocol

CH. Nagaraju, P. L. Mounika, K. Rohini, T. Naga Yaswanth, and A. Maheswar Reddy

Abstract The IIC protocol (inter-integrated circuit) is a communication bus protocol that allows many masters and slaves to communicate with each other. The I2C bus is a well-known and widely used communication technology in the electronics industry that can be found in a wide range of electronic devices. Circuit complexity is increasing as new technologies are created at a rapid pace. These circuits must be capable of self-testing. Built-in Self-Test (BIST) is a self-testing method that can be utilized instead of expensive testing equipment. The design and the creation of an Inter-Integrated Circuit (I2C) protocol that can self-test are presented in this work. The I2C uses the Verilog HDL language to achieve data transfer that is small, stable and reliable.

Keywords Verilog HDL · Built-In-Self-Test Architecture · Inter-integrated circuit

1 Introduction

1.1 Background

Circuit interconnected (I2C) protocols which enable devices that are speedier interact with a slower pace devices with no information misfortune. These are best interfaces in a system that involves a number of devices connected to together [1]. The development of System-on-Chip (SoC) makes the need for testing the interfaces of the embedded cores. Built-in-self-test (BIST) is an efficient solution to the requirement [2]. Philips semiconductors invented I2C in January 2000. An efficient and modern approach to design I2C with BIST stands for built-in self-test. It was proposed and modelled using Verilog HDL and focused on BIST, and its implementation is discussed here. BIST gives the predetermined testability necessities and best execution with least cost [3]. This framework is manufactured in a single chip and shows

CH. Nagaraju (✉) · P. L. Mounika · K. Rohini · T. Naga Yaswanth · A. Maheswar Reddy
Electronics and Communication Engineering, Annamacharya Institute of Technology and
Sciences, Rajampet, India
e-mail: chrajuaits@gmail.com

that the coding for this system is done using Verilog HDL, and design, testing and evaluation are done utilizing the Xilinx ISE tool. The framework requisites of bit error rate is low, high integration, and I2C can provide a low-cost solution.

In previous, the same BIST technology is implemented using the SPI protocol. The SPI protocol utilizes four wires for its data transfer from one master to the peripheral devices. For communication between a single master and one or more slaves, SPI stands for serial peripheral interface, generally known as the specification for synchronous serial interfaces utilized [4]. As the number of slaves grows, so does the complexity of the circuit; necessitating the development of a SPI has a self-testability capability modules in order to ensure that the circuits are fault-free. BIST stands for built-in self-test solution for circuit self-testing while also lowering maintenance and testing costs. This work introduces the BIST integrated SPI module design with a single-master and single-slave configuration, where the module exchanges 8-bit data, and the circuit under test (CUT) is self-tested for accuracy using the BIST feature [5]. This SPI module was created with the help of the EDA playground and Verilog hardware description language (HDL).

It necessitates more signal lines (wires) than other modes of communication. The communications must be well-defined in advance (we cannot transfer large amounts of data at any time). All communications must be controlled by the controller (peripherals cannot communicate directly with one another).

1.2 Objectives

To design Xilinx ISE Design Suite was used to simulate a BIST embedded I2C protocol with master–slave setup in Verilog HDL. The suggested model will make use of BIST's self-testability capability where the intended CUT employs the I2C. It also includes the ability to self-test to confirm that the circuit under test is working properly.

1.3 Need and Importance

- The I2C protocol (inter-integrated circuit) is a serial communication bus interface protocol. This specification is used to communicate between several masters and slaves.
- As the number of masters and slaves grows, so does the complexity of the circuit, necessitating the use of self-testability features to ensure circuits with no flaws.
- BIST stands for built-in self-test solution for circuit self-testing while also lowering maintenance and testing costs. It has been proposed to design a multiple master and slave configurations which are available in the BIST embedded I2C module.

2　Literature Review

The researchers in [6] proposed a functional verification plan for SPI protocol using Universal Verification Methodology (UVM)-based test bench architecture. The SPI model now has 100% functional coverage thanks to the careful development. Several test cases are admissible. The suggested UVM test bench creates random replies and compares them to the expected response using a scoreboard to ensure that it is working. This work's functional and cross-coverage report clearly shows that the full SPI operation is carried out appropriately, and the proposed method is well-suited to the verification of SoCs using the SPI protocol.

By building an IP core for SPI, it was established by the researchers in [7] as a paradigm for APB interaction. With a frequency of 16 MHz, SPI can transmit or receive data from a connected slave and APB-SPI controller and significant data that suggested work's flexibility. Model sim is used to simulate the planned work, and the GDI file for tape out is created using Quartus Lite 16. The modelled design work includes a simple interaction with extremely simple connections to the master's IO port that can be used to communicate with APB.

The authors in [8] provided an SPI architecture based on a single master and a single slave configuration utilizing independent slaves and a daisy chain setup of a single master and numerous slaves. The proposed SPI model was synthesized in Verilog HDL and validated in SystemVerilog before being implemented in Spartan 3E.

3　Methodology

3.1　Existing Methods

In existing methodology, we are frequently using SPI protocol as a means of communication used by a wide range of gadgets. One of the most significant advantages of SPI is that any amount of data can be delivered in an uninterrupted stream. The master is the device that assists in the generation of the primary. The SPI clock is used to sync the data. There can only be one SPI protocol master, although single to multiple slaves can be connected in a slave configuration. The link between a standard SPI master and SPI slave block is shown in Fig. 1. When multiple in a parallel arrangement, slave blocks are linked to the master block. Each slave block will thereafter have its own chip select signal for disconnecting from the master block, which will be amplified The SPI block's two data lines are master out slave in-signal from master to slave (MOSI) and master in slave out-signal from slave to slave (MISO) (master in slave out-signal from slave to master). It supports only one master and multiple slaves [9]. There is no acknowledgement in this type of communication from salve to master or from master to slave. This type of communication is used for short-distance communication.
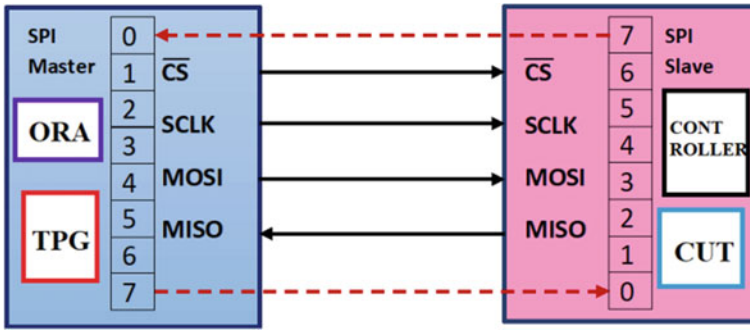
Fig. 1 BIST embedded SPI master–slave block

3.2　Proposed Method

SPI with BIST has been designed in prior work. BIST can be used to test itself in this fashion. Different blocks in the BIST scheme, the test controller, pattern generator, CUT and response analyser, are all included. To find the random number, the linear feedback shift register (LFSR) will be employed pattern creation in that BIST structure. The architecture for I2C with BIST mode as proposed is shown in Fig. 2. This mode of operation involves first the BIST in which the I2C tests itself, and second is the default setting. The device functions as the usual I2C device in this mode.

Normal mode is activated when reset is 1 and reset n is 0. BIST mode is enabled when reset is 0 and reset n is 1.
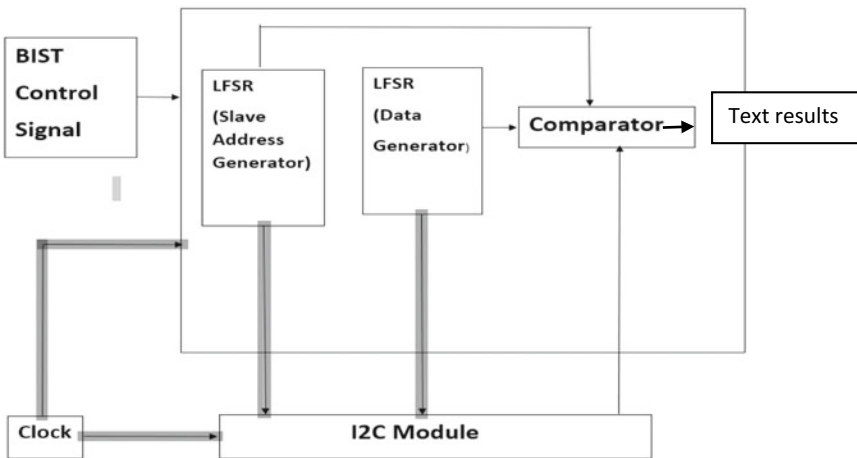
Fig. 2 Proposed block diagram
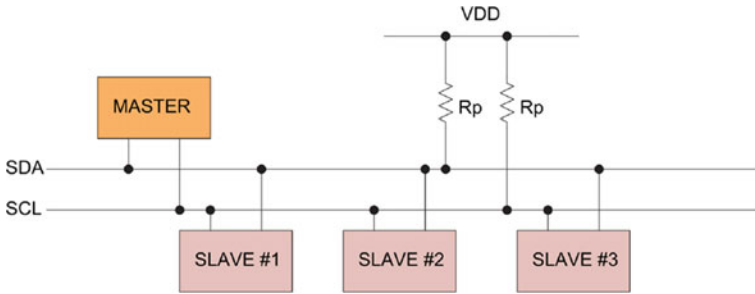
Fig. 3 I2C architecture

I2C Protocol Architecture

The I2C protocol architecture mainly consists of two wires: data bus and clock signal as shown in Fig. 3. The data transfer using this I2C protocol is bidirectional in nature which transfers data from master to slave or master's slave. The master produces a clock signal first to transfer the information to the slave, and then, the slave gives back acknowledgement signal, showing that the data transfer between master and slave involves the following components:

- **Slave Address**: This is the address of the slave to which the data is transferred from the master. It is 7-bit long.
- **Data Value**: The master transmits the information to the slave in the form of data value. Data values consist of 8 bits.

I2C Structure

Figure 4 depicts the operating modes of I2C and SPI. The two signals reset and reset _n are used to select the operation modes of I2C and SPI.

The data transmission begins when the master transmits a start bit to the master. When the clock signal swings from low to high, the data line swings from high to low, triggering the initial condition the slave's address sent. The slave recognizes the master, who then transmits the position of the register where the data is saved, which is followed by another acknowledgement from the slave to the master. The master transmits data to the slave via the data bus following the data transfer, the slave's master to the slave's master. Another acknowledgement is sent by the slave to the master. Finally, the master transmits the slave the stop bit, which is signalled by the
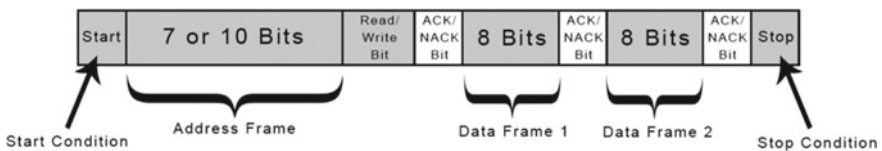
Fig. 4 I2C structure

Fig. 5 Schematic block
diagram
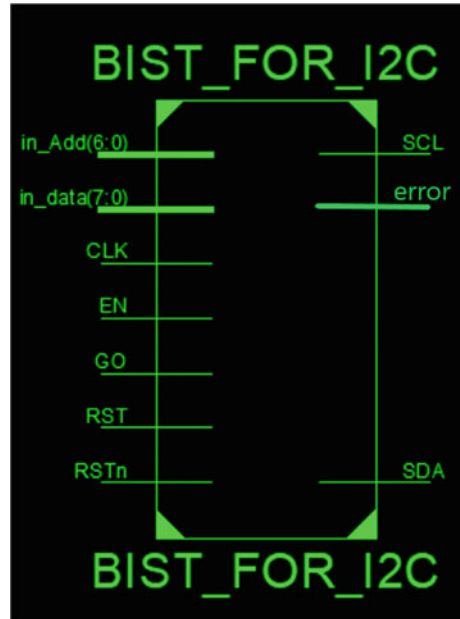
clock signal changing from low to high and the data line changing from low to high. Schematic block diagram of the work is shown in Fig. 5.

This hardware design is simpler when compared with existing methodology. This I2C protocol supports multiple masters and slaves. It uses only two wires for communication.

Implementation Tool

The tool used in this project is XILINX ISE tool in order to test the circuit. XILINX Integrated Synthesis Environment (ISE) is a synthesis environment. Xilinx programme is for HDL design synthesis and analysis. It allows programmers to synthesis designs, perform timing analysis, examine RTL diagrams, simulate the behaviour of a design to various stimuli and configure the device target.

4 Results

When RST is 1 and RSTn is 0 and EN is 1, this enables normal mode, and GO is used for resetting circuit as shown in Fig. 6. In_Add is input data address, and In_data is considered as input data in the system. The address and data are seen respectively in output addr and data. From counter we can see that data transmission is started, 3–9 address, 10 acknowledgement, 11 read–write, 12–19 data, 20 acknowledgement, and 21 is stop bit when both scl and sda are high.
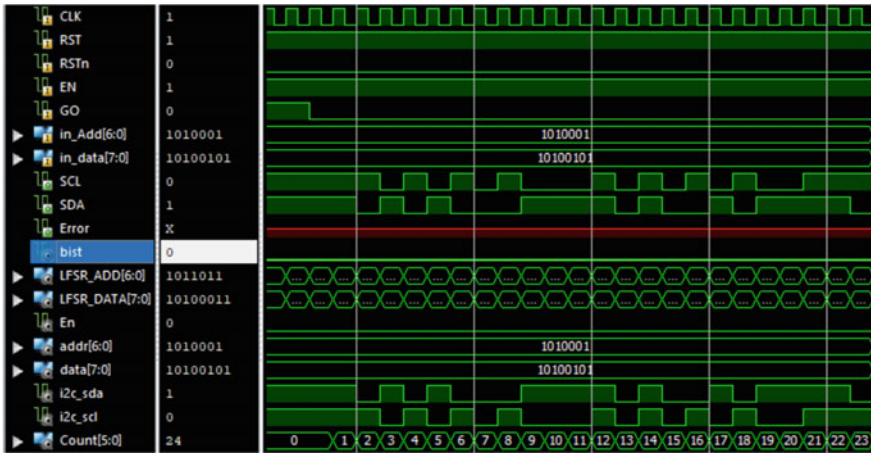
Fig. 6 Normal mode

When RST is 0 and RSTn is 1 and EN is 0, this enables BIST mode as shown in Fig. 7. GO is used for resetting circuit. In_Add is input data address and Data is input data which does not show up in output. LFSR_ADD (LFSR address) and LFSR_DATA (LFSR data) are seen in output address and data, respectively. From counter, we can see From 2 data transmission is started, 3–9 address, 10 acknowledgement, 11 read write, 12–19 data, 20 acknowledgement, and 21 is stop bit when both scl and sda are high.
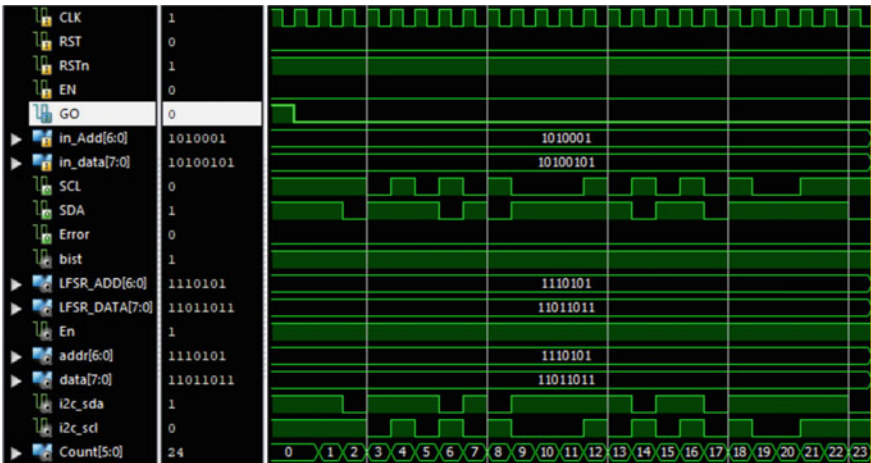
Fig. 7 BIST mode

Advantages

- It is faster than asynchronous serial.
- It supports multiple slaves or peripherals.
- It supports full-duplex communication.

Limitations

- It employs one CS line per slave; therefore, if there are a lot of slave devices in the design, the hardware complexity rises.
- In order to add a slave device, we will need to alter the programme and add an extra CS line.
- In SPI, there is no acknowledgement.
- More number of wires is required in this of communication methods.
- It supports only single master and does not support multiple masters.

5 Conclusion

The implementation of I2C with BIST capabilities is provided in this work. Verilog HDL is used to develop and simulate all of the modules. In comparison with standard I2C, this one is substantially less expensive, faster and more stable. This I2C control bus architecture can enable processes where only one switch is pressed to test the system. As a result, testing this protocol bus will take less time and cost less money.

A Xilinx-based I2C implementation with BIST functionality is provided in this work. Here, all the modules are Verilog HDL which were used to design and simulate the system. The system will then be activated. The I2C bus is a simple two-wire serial bus which saves time and money. As a result, ICs contain fewer pins and there are fewer interconnections. PCBs are smaller and less expensive. Interconnected circuits are more reliable and speedier and far more adaptable. It refers to a way by which a circuit can test itself. I2C with BIST capability is possible to provide further benefits. The simulation produces a proper outcome. It also saves time and effort and significantly reduces the cost of circuit testing.

Bibliography

1. I2C Bus Specification (2000) Version 2.1, Philips Semiconductors
2. Venkateswaran P, Mukherjee M, Sanyal A, Das S, Nandi R (2009) Design and implementation of FPGA based interface model for scalefree network using I2C bus protocol on Quartus II 6.0. In: Proceedings of 4th international conference on computers and devices for communication, pp 1–4
3. Singh R, Sharma N (2013) Prototyping of on-chip I2C module for FPGA Spartan 3A series using Verilog. Int J Comput Appl 68(16)
4. Bushnell M, Agarwal VD (2000) Essentials of electronic testing for digital, memory and mixed-signal VLSI circuits. Kluwer Academic Publishers

5. Jamuna S, Agrawal VK (2011) Implementation of BIST structure using VHDL for VLSI circuits. Int J Eng Sci Technol 3(6):5041–5048
6. Vineeth B, Tripura Sundari BB (2018) UVM based testbench architecture for coverage driven functional verification of SPI protocol. In: 2018 International conference on advances in computing, communications and informatics (ICACCI), pp 307–310. https://doi.org/10.1109/ICACCI.2018.8554919
7. Hafeez M, Saparon A (2019) IP Core of Serial Peripheral Interface (SPI) with AMBA APB Interface. In: 2019 IEEE 9th symposium on computer applications & industrial electronics (ISCAIE), pp 55–59. https://doi.org/10.1109/ISCAIE.2019.8743871
8. Pallavi Polsani V, Priyanka B, Padma Sai Y (2020) Design & Verification of Serial Peripheral Interface (SPI) Protocol. Int J Recent Technol Eng (IJRTE) 8(6)
9. Saha S, Rahman MA, Thakur A (2013) Design and Implementation of a BIST embedded high speed RS-422 utilized UART over FPGA. In: Proceedings of 4th IEEE international conference on computing, communication and networking technologies (ICCCNT)

Implementation of Finite Element Method in Opera Software for Design and Analysis of High-Voltage Busbar Supports

O. Hemakesavulu, M. Padma Lalitha, and N. Sivarami Reddy

Abstract Busbar supports are an important part of the transmission line, as they are utilized to carry conductors using cross-arms and insulators. Busbar supports must be able to bear the busbar's mechanical forces and tension. The transmission tower is another name for the busbar support. This research report discusses the effort to construct a monopole support structure for a 33 kV transmission line. The proposed research was carried out utilizing the Opera-electromagnetic field. This analysis program was used to create a three-dimensional model, which was then simulated using the FEM. The field characteristics, including the electric field intensity distribution and potential distribution, are returned by the Opera-3D-TOSCA and lossy dielectric solvers.

1 Introduction

High-voltage busbar supports, often known as towers, are used in transmission lines to transport conductors over long distances. These busbars are connected to the cross-arms using insulator banks. Depending on the voltage rating line, the support structure varies. The voltage and electric potential near conductors are four to six times higher than elsewhere on the surface. Several experimental and calculation methodologies for calculating the voltage distribution have been established [1, 2].

Numerous investigational and modeling techniques related to determination of the potentials and electric fields have been developed in [3–6]. Researchers can investigate the models' behavior with extraordinarily intricate geometry by means

O. Hemakesavulu (✉) · M. Padma Lalitha · N. Sivarami Reddy
Professor, Department of Electrical and Electronics Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, Andhra Pradesh, India
e-mail: hkesavulu6@gmail.com

N. Sivarami Reddy
Dean R&D, Professor, Mechanical Department, Annamacharya Institute of Technology & Sciences, Rajampet, Andhra Pradesh, India

of simulation methods rather than analytical or experimental methods. Many articles have suggested that the voltage distribution and electric field with asymmetric boundary conditions be calculated using the finite element method (FEM) [3–5, 7].

The potential and electric field around and inside the insulator string were replicated in Opera using a 3D FEM model in [6] and [7]. Numerical algorithms have been developed to determine the potential and electric field within and around HV equipment rather of relying on analytical or experimental methods.

Faisal [6], investigated the potential and electric field of a polluted insulator using boundary integral equations. COBHAM [8], describes how to use the Opera software, create a model, and mesh it.

2   Technical Specifications of 33 kV Busbar Are of Use

The support structure must be able to carry the line conductors, as well as the appropriate insulators, fittings, and equipment, under the stipulated conditions. Both 100 and 232 mm^2, conductors are used for 33 kV transmission lines. Intermediate pole constructions for conductor sizes of 100/232 mm^2 must be double pole or single pole with pin or post insulators. The usual span length for 100 mm^2 should be 80–100 m and 45–55 m for 232 mm^2. The busbar nowadays used are of ACSR type and AAAC type. The minimum clearance between the ground and line conductors or other objects is represented in Table 1.

IS: 5613 specifies the wind pressures that must be applied to the insulators, conductors, and support structures for 33 kV lines. The working load on the structures should correspond to the kind of loads that are expected to be applied to them during the course of their service life.

The safety factor for various types of support systems varies based on the structure type, and for mechanically processed concrete supports, it is 2.5.

Table 1 Minimum ground clearance in different situations

Situation	Minimum clearances (Mts)
Above open country	5.22
Above the roadway	6.20
By the side of roadway	5.80
Over telecommunication channel	2.54
Along the river	6.20
Above the railway line	14.15
Along 11 kV or low-voltage line	2.43
Underneath high-voltage line	2.45

3 Finite Element Method

Due to its high capacity to address the regions of extremely complicated geometries, this approach is becoming progressively more popular. It splits the investigated region into a number of sub-domains called elements (typically triangles, quadrilaterals, tetrahedral, or hexahedral).

Because the field on every element is modeled by a simple algebraic expression, the field values on a small collection of nodal points or edges are found as the solution of a linear set of equations. When dealing with unbounded geometry, there are a number of options. While boundary elements are capable of accurately depicting an infinite region, differential equation solvers can acquire a good approximation by enclosing the region of interest in a big box. The validity of the solution is influenced by the size of the constraining box (or outside boundary), as well as the number and distribution of elements.

4 Simulation Method

Opera-3D [8] is a suite of three-dimensional electromagnetic field analysis programs. Using the finite element method, the application solves partial differential equations (Poisson's, Helmholtz, and diffusion equations) that explain field behavior. Field solution modules for constructing electric insulating components employing conducting dielectric materials in transient and steady-state circumstances are included in the lossy dielectric solver. The program solves a current flow problem before applying the results to an electrostatic problem. By solving the conduction (current flow) equation, the software calculates the potential. The intensity of the electric field E is given by

$$\boldsymbol{E} = -\nabla V \tag{1}$$

The electric flux density divergence, $\mathbf{D}$, is interrelated to charge density ρ:

$$\nabla \boldsymbol{D} = \rho \tag{2}$$

By the combination of Eqs. (1) and (2) and establishing the permittivity, dielectric tensor ε ($\mathbf{D} = \varepsilon \mathbf{E}$) occurs the standard Poisson's equation description of the electrostatic potential:

$$\nabla \varepsilon \nabla V = -\rho \tag{3}$$

A analogous equation arises for current flow problems,

$$\nabla \sigma \nabla V = 0 \tag{4}$$

where σ is the conductivity, and $\mathbf{J} = \sigma\mathbf{E}$.

The electric and magnetic fluxes are generally treated in a combined way in time-varying situations. The software solves for assuming that inductive effects are insignificant in semi conducting dielectric situations.

$$\nabla\varepsilon_c\nabla V = 0 \tag{5}$$

The model is analyzed using FEM, which divides the model's region into linear triangle parts. The density of the finite element mesh is set to be higher in important portions of the insulator than in the remainder of the model. This provides the user with two benefits: reduced complexity and reduced simulation time.

5 Design of High-Voltage Busbar Supports in Opera

In the present work, Monopole tower structure is designed for 33 kV transmission line and analyzed using FEM in Opera software. Here, model is designed as per the parameters received from industrial experts. To reduce the complexity of the design, the cross-arms are neglected, and the transmission line is represented as block of plate type in which potential is applied using boundary conditions. In this work, we analyzed the potential distribution and electric field distribution in and around the tower structure. Here, a Cartesian patch is used to observe the fields around the surroundings. Here, only quarter of the model is analyzed using model symmetry. Each busbar support is furnished with insulating plates (known as sheds or skirts) placed around the device primarily to increase the length of the creep path for fields that track down the side of it. Figure 1 represents the Monopole tower model designed in Opera.

The voltage and electric field distribution around the tower were calculated using the Opera-3D TOSCA (electrostatic solver) in combination with the lossy dielectric solver. The simulated model is shown in Fig. 2.

When simulated, this model has 27 cells, 204 faces, 390 edges, and 240 vertices, resulting in 175,730 elements and 107,691 nodes. The simulation of 67,166 equations took 3 min and 25 s. This model is solved by considering the conductivity only. The potential distribution around the model is taken on a Cartesian patch and is shown in Fig. 3.

Figure 4 gives us distribution of electric field in and around the tower and surrounding air. Electric stress is greatest around metal tip and conductors and gradually decreases as it moves far away with respect to the conductor. The electric field parameters along a line that passes through the sheds can be viewed by drawing a Y-directed line from coordinate (7, 0, 0) to (7, 60, 0) with at least 240 steps (Fig. 5).
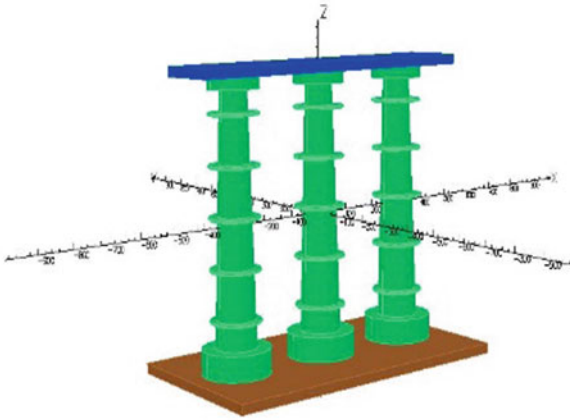
Fig. 1 Monopole tower model of 33 kV
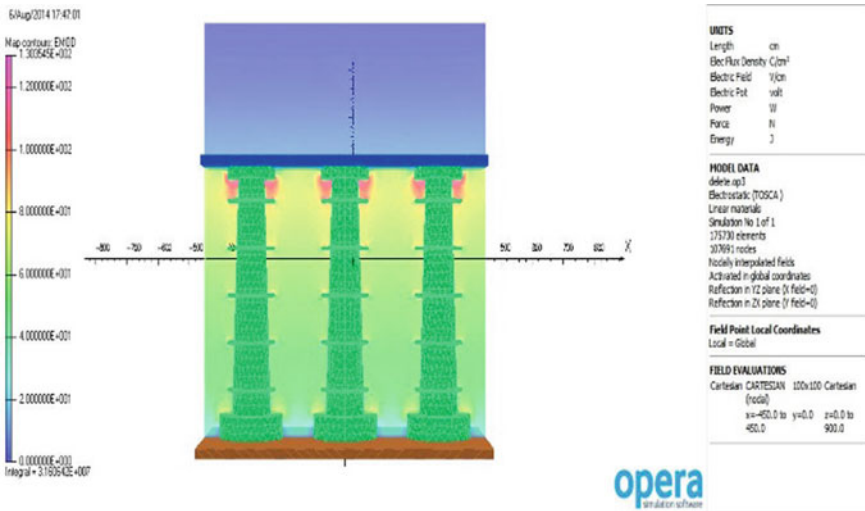
Fig. 2 Simulated model

6 Conclusion

In the present work, design of a monopole tower model for 33 kV busbar supports using Opera is focused. The concentrations of elements are high near the critical regions and are less at the outer boundary regions where ever the importance is low. The results show that the concentration of potential and electric field is much

Fig. 3 Potential distribution around the tower on a Cartesian patch

Fig. 4 Electric field distribution around the tower on a Cartesian patch

Component: EMOD, Integral = 33104.1166738253

Fig. 5 Electric field counters along a *Y*-directed line

higher at the busbar than at other locations on the surface, and that these distributions similarly diminish as one moves away from the surface. The electric field stress near the busbar grows as the conductivity of the busbar increases.

References

1. Dhalaan JMA, Elhirbawy MA (2003) Simulation of voltage distribution calculation methods over a string of suspension insulators. Transmission and distribution conference and exposition, 2003 IEEE PES, vol 3, 7–12 Sept 2003, pp 909–914
2. Farag ASA, Zedan FM, Cheng TC (1993) Analytical studies of HV insulators in Saudi Arabia-theoretical aspects. IEEE Trans Electr Insul 28(3):379–391
3. Kontargyri VT, Gonos IF, Stathopulos IA, Michaelidis AM (2003) Calculation of the electric field on an insulator using the finite elements method. In: 38th international Universities power engineering conference (UPEC 2003), Thessaloniki, Greece, 1–3 Sept 2003, pp 65–68
4. Kontargyri VT, Gonos IF, Stathopoulos IA, Michaelides AM (2006) Measurement and verification of the voltage distribution on high voltage insulators. In: 12th biennial IEEE conference on electromagnetic field computation, 1–3 May 2006, pp 326–326
5. Kontargyri VT, Gonos IF, Ilia NC, Stathopulos IA, Simulation of the electric field on composite insulators using the finite element method

6. Faisal ShM (2009) Simulation of the electric field distribution on ceramic insulator using finite element method. Eur Trans Electr Power 19(3):526–531
7. Rasolonjanahary JL, Krähenbühl L, Nicolas A (1992) Computation of electric fields and potential on polluted insulators using a boundary element method. IEEE Trans Magn 38(2):1473–1476
8. COBHAM, OPERA-3d user guide—16R1. Cobham technical services. Vector Fields Software, UK, 1999–2013

InterCloud: Utility-Oriented Federation of Cloud Computing Environments Through Different Application Services

Rajesh Tiwari, Rajeev Shrivastava, Santosh Kumar Vishwakarma, Sanjay Kumar Suman, and Sheo Kumar

Abstract This white paper presents the vision, challenges, and key components of InterCloud for value-driven organizations in cloud computing environment. The related cloud environments support application scaling and career clouds. This method has been validated by performing a series of rigorous performance assessments using the Cloud Sim toolkit. The diagram shows that the hybrid cloud plan has great potential as it provides remarkably fast response times and resource savings in high-volume scenarios.

Keywords InterCloud · Cloud environment · Cloud security · Federated environment

1 Introduction

Leonard Kleinrock, one of the principal analysts for the first ARPANET (Progressed Investigate Ventures Organization) project that laid the foundations of the Internet in 1969, said [1]: When you grow up and modernize, you will see a proliferation of "computer utilities" that will benefit people at home and at work across the country, such as electricity and telephone utilities. This vision of computing utilities, based on a benefit model, envisioned a massive transformation of the entire computing industry

R. Tiwari · S. Kumar
CMR Engineering College, Hyderabad, Telangana, India

R. Shrivastava
Princeton Institute of Engineering and Technology for Women, Hyderabad, India

S. K. Vishwakarma (✉)
Manipal University Jaipur, Jaipur, India
e-mail: santosh.kumar@jaipur.manipal.edu

S. K. Suman
St. Martin's Engineering College, Hyderabad, Telangana, India

in the twenty-first century, in which computing services are delivered quickly, on-demand and on-demand, just like any other managed service available in today's society. Basically, the computer service customer (the buyer) has to pay the provider as if they were dealing with computer maintenance. As customers grow, they no longer have to make significant contributions or struggle to build and maintain complex IT infrastructures. In these examples, clients approach the organization on an as-needed basis, regardless of where the company is promoted. This is called Hands-on Computing, and more recently, Cloud Computing [2].

The last term mentioned refers to systems as "the cloud" that allow companies and customers to access their application organization from anywhere in the world if needed. Thus, cloud computing in general can be classified as a high-level thinking for rapidly delivering IT organizations powered by modern data centers containing an organized set of virtual machines. Cloud computing is moving toward creation, organization, and computer programs (applications) by organizations, opening up as subscription-based organizations in the form of paid subscriptions for buyers. These organizations are referred to in the industry alone as "Systems as a Benefits" (IaaS), "Organizations as Benefits" (PaaS), and "Programs as Benefits" (SaaS). According to a February 2009 Berkeley report, "Cloud computing, the long-standing dream of computing as a utility has the potential to transform the broader IT industry by making software as a service more attractive" [3].

Clouds were designed to manage data centers in the past, creating them as a set of virtual organizations (hardware, databases, user interfaces, how applications are handled) so that customers can request and submit applications on-demand in a competitive manner. The cost depends on the client's requirements for Quality of Advantage (QoS) [2]. Engineers with creative minds for cutting-edge web organizations do not need huge capital expenditures on equipment to send services or recruit people to operate them [3]. This gives IT companies a significant advantage. This frees them from the burden of setting up critical hardware (servers) and computer software systems to focus more on growing and sharing the organization. The commercial potential of cloud computing is assessed by several large companies including IDC, and organizations around the world have contributing to cloud computing will receive $42 billion in 2012 from $16 billion by 2008. Additionally, various applications using utility-oriented computing systems such as the cloud are created primarily to facilitate or catalyze manufacturers that unite buyers and sellers. This brings trillions of dollars to the utility/ubiquitous computing industry, popularized by the Charge Enchant [4], co-founder of Sun Microsystems. "It will take some time for such a market to form," he said. "It would be foolish to predict directly which companies will get attention. Many of them haven't really been made yet."

2   System Architecture

Figure 1 shows the high-level components of a service-oriented building system, including the management of brokers and client facilitators that support cloud service alliances such as application planning, asset allocation, and workload movement. Provides cloud capabilities as part of a single-asset leasing solution. This platform will facilitate the consolidation of cross-domain capabilities for adaptive, reliable, and energy-efficient access to on-demand environments based on new innovations in virtualization [5, 6].

Cloud Exchange (CEx) acts as an advertisement producer, bringing makers and clients together. Coordinated application brokers' regulation necessities and benchmark them against the open offerings as of now sent by cloud brokers. It underpins commerce in cloud organizations based on competitive trade models [7] such as product and bargain markets. CEx issues licenses to individuals (cloud resellers and cloud brokers) to find vendors and customers offering suitable products.

These marketplaces pave the way for organizations to engage their products and create passionate trading platforms based on service level agreements. The SLA discloses the non-infringing component of the benefit to be delivered based on an estimate agreed by all parties. The driving force to exclusively capture and meet the requirements, and the principles. With an internal cash storage system they facilitate, SLA-related financial transactions between individuals take place in a safe and secure environment. Each client in the combined organization must create a Cloud Brokering
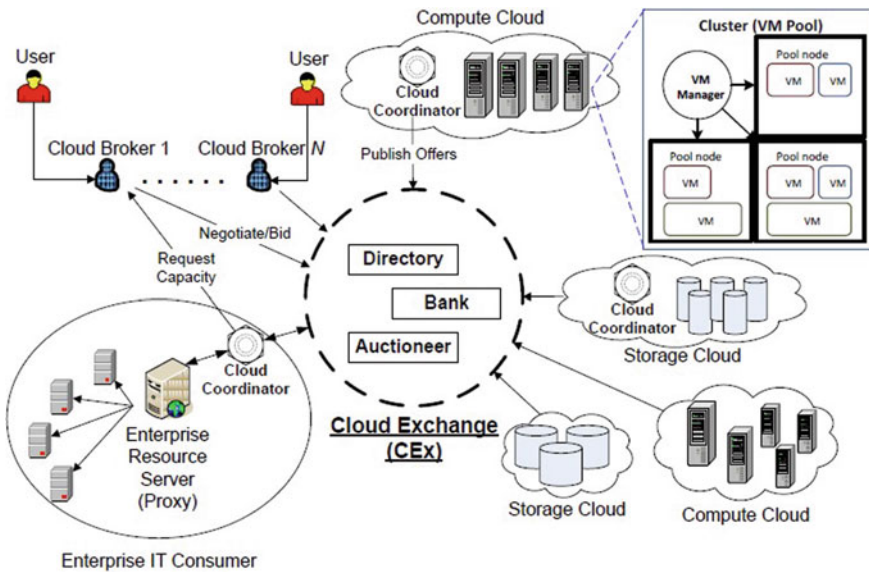
Fig. 1   Federated network of clouds mediated by a cloud exchange

Benefit that enables them to efficiently negotiate profitable contracts with Cloud Facilitators through the trading opportunities they discover in Cloud Trade.

3   Elements of InterCloud

3.1   *Cloud Coordinator (CC)*

The Cloud Facilitator Benefit is designed to manage specific space clouds and register them in a common alliance based on exchanges and trading market contracts. Provides programming, management, and hosting environments for cloud league applications.

Scheduling and Allocation. Virtual machines in the cloud center are assigned based on your QoS goals and your organization's goals in the cloud. When a client application is received, the scheduler does the following: It (i) instructs the application build engine to open roughly the software and hardware organization needed to fulfill the request locally [8], (ii) requests to provide feedback on the sensor component and usage status of neighboring key cloud nodes, and (iii) request a tribune drive and action close to the obligations of the submitted request.

Market and Policy Engine. The SLA module stores the cloud benefit terms for each Cloud Broker user for each customer. These terms and conditions allow the scoring mechanism to choose how to claim benefits based on public offering and required cloud computing resource requirements. The accounting module stores information about the actual resource usage for each request so that the overall utilization for each client can be calculated. At this time, the charging module charges the client within the same way.

Application Composition Engine. This Cloud Facilitator component includes the ability to interact on request with a database backend, such as SQL data administration given by Microsoft Purple Blue, an application server like web information that empowers application engineers to construct and convey applications. The reservation work has been actuated. Server IIS has a secure ASP.Net scripting motor for web applications and a Cleanser-based web administration API for automatic interaction and integration with other applications and information.

3.2   *Cloud Broker (CB)*

The cloud broker on behalf of the customer recognizes the appropriate cloud benefit providers through the cloud exchange and organizes the resource operation through the cloud facilitator to meet the QoS requirements of the customer.

User Interface. The cloud broker recognizes the relevant cloud computing benefit providers through the cloud exchange on behalf of the customer and works with the cloud broker to coordinate resource operations to meet the customer's quality of service requirements. It provides communication between the client application interface and the broker. The application interpreter decodes the client application's execution prerequisites that bind the action to be performed. It has errors and additional validation of input images, data records (if required), and near-false information and outputs (if indicated) and the required quality of service.

Core Services. They provide the best broker experience possible. Advantage Authority offers trading for cloud organizations on Cloud Trade. The planner selects the best cloud organization for the client application based on the application's needs and benefits. The benefits screen periodically checks the availability of known cloud organizations, finds open unused ones, and maintains the status of cloud organizations.

4 The Performance of the Federated Cloud Environment

The first attempt shows that the integrated cloud computing environment has the potential of convey superior execution and benefit quality compared to existing non-coherent approaches. To this conclusion, a tournament modeling reverted environment of three Cloud providers and one customer (Cloud Broker) is modeled. Each vendor initiates the Sensor component, which can robustly detect proximity-related accessibility data. In addition, the discovered metrics are detailed for the Cloud Facilitator to use the data to make load migration selections.

We are evaluating a load migration action plan that explicitly implements the online migration of VMs between sites that are bound together. This is because the cloud environment does not have the free space, required for the virtual machine from the initial provider. The move phase includes the following steps: (i) generate virtual machine events with the same action plan supported by the target provider [9, 10]; and (ii) move applications assigned to the initial VM from the target provider to the most recently created VM. The combined cloud provider architecture is based on the topology shown [11–14] in Fig. 2.

5 Conclusion

Improvement of principal strategies and program frameworks that coordinated conveyed clouds in a combined mod is basic to empowering composition and sending of versatile application administrations. We accept that results of this investigate vision will make noteworthy logical headway in understanding the hypothetical and viable issues of building administrations for combined situations. The coming about
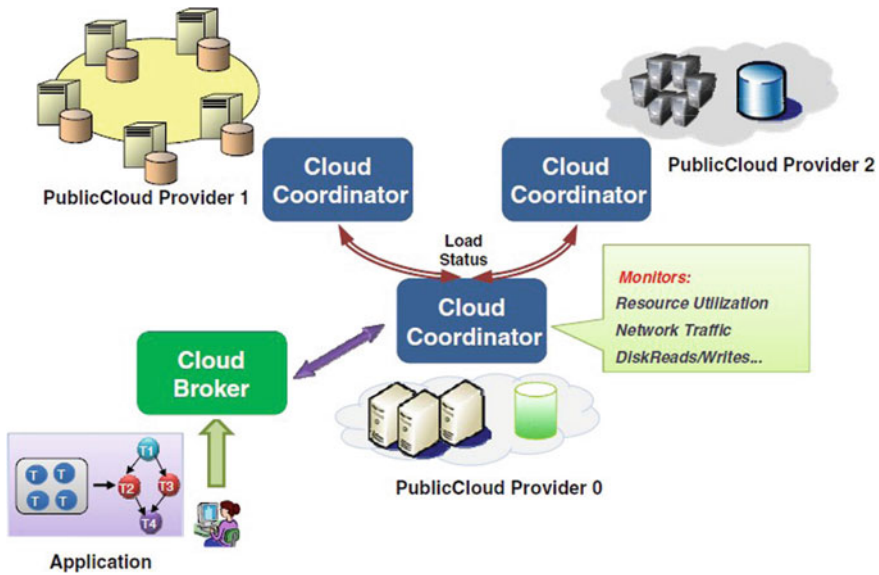
Fig. 2 Network topology of federated data centers

system encourages the unified administration of framework components and secures clients with ensured quality of administrations in expansive, unified and exceedingly energetic situations. The components of the proposed system offer effective capabilities to address both administrations and assets administration, but their end-to-end combination points to significantly make strides the viable utilization, administration, and organization of Cloud frameworks. This will deliver updated degrees of versatility, flexibility, and straightforwardness for organization and transport of organizations in collusion cloud. Future work will focus on an integrated approach to delivery and transfer to organizations in relevant contexts.

References

1. Kleinrock L (2005) A vision for the internet. ST J Res 2(1):4–5
2. Buyya R, Yeo C, Venugopal S, Broberg J, Br&ic I (2009) Cloud computing & emerging IT platforms: vision, hype, & reality for delivering computing as the 5th utility. Future Gener Comput Syst 25(6):599–616
3. Armbrust M, Fox A, Griffith R, Joseph A, Katz R, Konwinski A, Lee G, Patterson D, Rabkin A, Stoica I, Zaharia M (2009) Above the clouds: a Berkeley view of cloud computing. University of California at Berkley, USA. Technical Rep UCB/EECS-2009–28
4. London S (2002) Inside track: the high-tech rebels. Financial Times
5. VMware: Migrate virtual machines with zero downtime
6. Barham P et al (2003) Xen & the art of virtualization. In: Proceedings of the 19th ACM symposium on operating systems principles. ACM Press, New York

7. Buyya R, Abramson D, Giddy J, Stockinger H (2002) Economic models for resource management & scheduling in grid computing. Concurrency Comput Pract Experience 14(13–15):1507–1542
8. Weiss A (2007) Computing in the clouds. NetWorker 11(4):16–25
9. Buyya R, Abramson D, Venugopal S (2005) The grid economy. Special issue on grid computing. In: Parashar M, Lee C (eds) Proceedings of the IEEE, vol 93(3), pp 698–714. IEEE Press, Los Alamitos
10. Yeo C, Buyya R (2006) Managing risk of inaccurate runtime estimates for deadline constrained job admission control in clusters. In: Proceedings of the 35th international conference on parallel processing, Columbus, Ohio, USA
11. Sulistio A, Kim K, Buyya R (2008) Managing cancellations & no-shows of reservations with overbooking to increase resource revenue. In: Proceedings of the 8th IEEE international symposium on cluster computing & the grid, Lyon, France
12. Chu X, Buyya R (2007) Service oriented sensor web. In: Mahalik NP (ed) Sensor network & configuration: fundamentals, standards, platforms, & applications. Springer, Berlin
13. Awantika PM, Tiwari R (2020) A novel based AI approach for real time driver drowsiness identification system using Viola Jones algorithm in MATLAB platform. Solid State Technol 63(05):3293–3303. ISSN: 0038-111X
14. Shrivastava R, Singh M, Thakur R, Subrahmanya Ravi Teja KS (2021) A real-time implementation for the speech steganography using short-time Fourier transformer secured mobile communication. J Phys Conf Series 2089:012066. https://doi.org/10.1088/1742-6596/2089/1/012066

Cross-Cultural Translation Studies in the Context of Artificial Intelligence: Challenges and Strategies

Rajeev Shrivastava, Mallika Jain, Santosh Kumar Vishwakarma, L. Bhagyalakshmi, and Rajesh Tiwari

Abstract It is an enormous challenges and strategies in cross-cultural. As a result of the analysis in artificial intelligent, it isn't a primary language for analysis teams. In society analysis, translation provides an additional challenge and strategy. In society analysis, the interpretation and comprehension of which suggest of data is concentrate. The aim of this text is to supply an overview of the interpretation methodology and explore variety of the challenges, like difficulties notice associate degree applicable translator, and also the importance of communication between the investigator and also the translator. A developing commercial center for dialect benefit on and rapidly developing innovation has introduced in an awfully unused wave of T&I computer program bundle advancement. In spite of interpretation tools' wide choice of applications conjointly the current blast of AI, interpretation memory and interpretation tools' machine learning calculations are expelled from palatable in giving dialect arrangements. Tragically, relate degree overreliance in this innovation. By dissecting cases wherever computational phonetics gave erroneous recommendations, this content points to investigate the association between rising innovation and antiquated expertness in T&I. The author believes that innovation and expertness aren't in an awfully zero-sum amusement, and innovation doesn't deny interpreters of work openings, in any case or maybe their expository aptitudes, fundamental considering, and stylish interests.

R. Shrivastava
Princeton Institute of Engineering and Technology for Women, Hyderabad, India

M. Jain
Kalaniketan Polytechnic College, Jabalpur, India

S. K. Vishwakarma (✉)
Manipal University Jaipur, Jaipur, India
e-mail: santosh.kumar@jaipur.manipal.edu

L. Bhagyalakshmi
Rajalakshmi Engineering College, Chennai, Tamil Nadu, India

R. Tiwari
CMR Engineering College, Hyderabad, Telangana, India

1   Introduction

It's the science and building of building machines to illustrate insights especially
seeing, discourse acknowledgment, decision-making, and interpretation between
dialects like identities. AI is that the reenactment of human insights forms by
machines, especially pc frameworks. This incorporates learning, thinking, arranging,
self-correction, drawback assurance, information outline, recognition, movement,
control, and inventiveness. it's a science and a collection of machine procedures that
unit galvanized by the way inside which identities utilize their framework anxious
and their body to feel, learn, reason, and act [1]. AI is clarified to machine learning
and profound learning whereby machine learning makes utilize of calculations to
induce designs and create experiences from the data they're locked in on. Profound
learning may be a set of machine learning, one that brings AI closer to the objective
of facultative machines to assume and work as human as potential [2]. AI may be
far from being obviously true subject and is commonly portrayed in an exceedingly
negative way; a few would choice it a favoring in mask for businesses, while for a
number of it's an innovation that imperils the unimportant presence of human race
since it is without a doubt able of seizing and ruling soul, in any case really computing
has influenced our way either specifically or in a roundabout way and forming the
longer term of tomorrow. AI has as of now become AN intrinsic a portion of our
way of life and has significantly wedged our way in spite of the basic employments
of computerized collaborators of versatile phones, driver assistance frameworks, the
bots, writings and discourse interpreters, and frameworks that help in suggesting
stock and administrations and customized learning. Each rising innovation may be
a supply of each eagerness and skepticism [3]. AI may be a supply of each benefits
and downsides in various sees. In any case, we need to defeat bound challenges some
time recently we'll take note truth potential and tremendous transformational capa-
bilities of this rising innovation, a number of the challenges related with computing
are shown in Fig. 1.

1.1   *Challenges*

Following are the challenges in AI. The use case template of AI is shown in Fig. 2.

Building Trust. The AI is all concerning science, innovation, and calculations that
to a great extent people unit of measurement uninformed. It trusts for them.
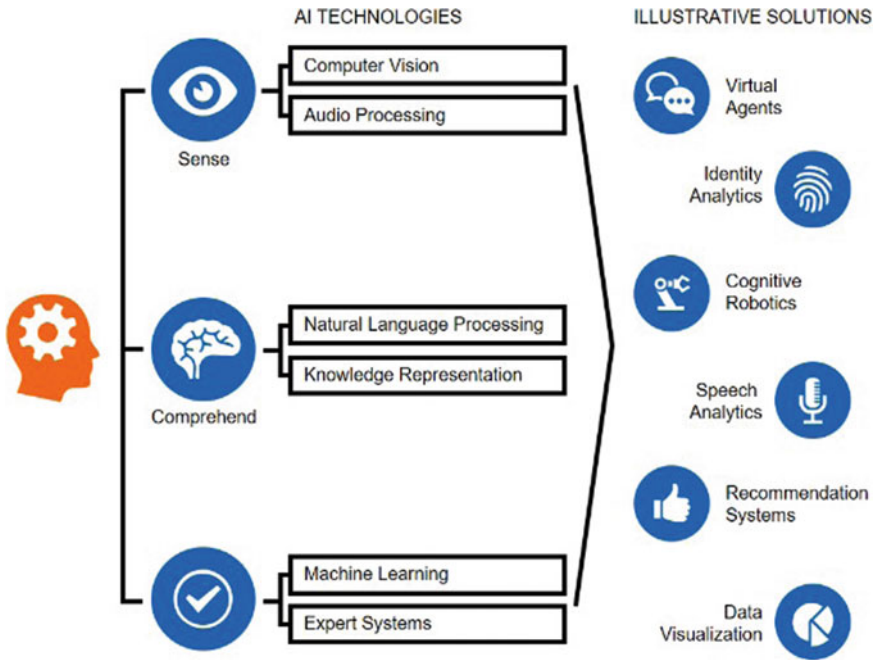
Fig. 1 AI technologies
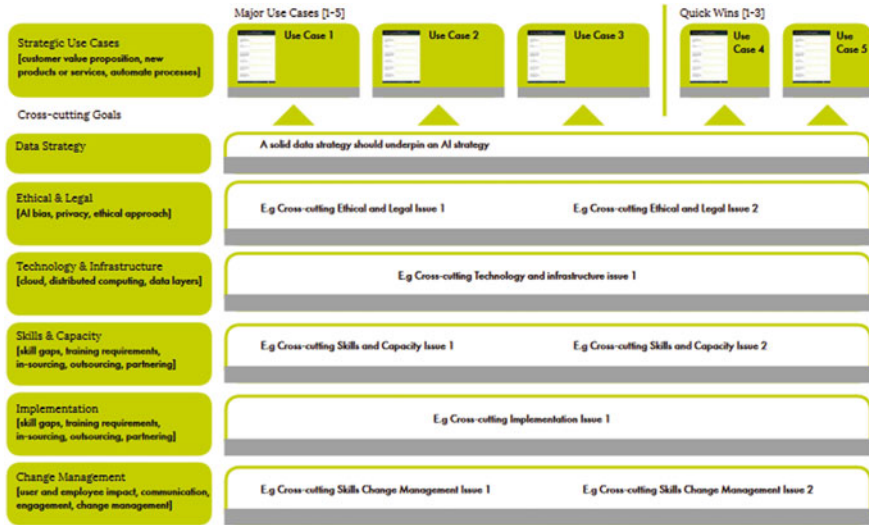
Fig. 2 AI use case template

AI Human Interface. There's an expansive brief of working faculty having information analytics and information science abilities. Those progressively may be deputed to initiate most yield from computer science.

Investment. AI innovation at scale and so been dispossessed of cost great thing almost scale. Once decades of hypothesis and passable uneasiness concerning the social suggestions of gathering and without a doubt de-stabilizing AI innovation for gathering and recording machine downside, AI investor's unit bit doubtful from stopping their cash in potential new companies.

Software Malfunction. Counterfeit insights choose and capacity is mechanically ceded to drive code recording machine apparatus. Additionally, since the deficiency of capacity of kinsfolk to discover out and see be that as it may these instruments work, they require exceptionally small or no administration over the framework that's extra troublesome as driven machine frameworks. It gets to be extra winning and advanced.

Non-invincible. (Can supplant as it were beyond any doubt assignments) like a few diverse innovations. AI furthermore possesses restrictions. It simply cannot supplant all errands; In any case, it will lead to rising modern work space with completely diverse quality work profile.

1.2 Strategies

Begin by completing the AI Utilize Case Format for each of your AI priorities/projects. AI utilize cases will be distinctive for each company and will be driven by your vital goals. Be that as it may, a few common ways to utilize AI include:

- Developing a parcel of cleverly item
- Developing a parcel of shrewdly administrations
- Creating trade forms more brilliant
- Automating tedious trade errands
- Automating creating processes.

Data Strategy. The AI methodology ought to be supported by a radical, up-to-date data procedure. After all, AI essentially doesn't work whereas not data. So, in case it's been a brief, whereas since you checked out your data methodology, now's a better than average time to come back it. AI needs will affect or revision beyond any doubt ranges of this data strategy.

Ethical and Legal Problems. There are tons of ethical and lawful contemplations around AI, and you're likely to come across assortment of identical issues for each utilize case. This opportunity is to recognize those crosscutting subjects. For case, no matter approach you utilize AI, consent and data security are key concerns. AI is liberated from inclination and segregation, which the approach you're abuse AI is

ethical, AI have to be utilized for the awesome of the commerce, its specialists, and its customers.

Technology and Infrastructure. In this segment, you must point to spot common topics around innovation and foundation. So, what innovation necessities and challenges unit a proportionate over your various AI utilize cases? At this organize, I take note it makes a difference to think almost the four layers of data and decide what innovation you'll need for each layer:

- Grouping data
- Storing data
- Process (examining) data
- Communication bits of knowledge from information.

Skills and Capacity. There is a gigantic abilities hole in AI and data. It's so greatly without a doubt that your commerce can have various aptitudes holes to plug, which these aptitudes holes could be common over the different utilize cases. There may, as a case, be crosscutting coaching necessities. Or possibly you'll be compelled to lease modern representatives or accomplice with relate degree outside AI provider.

Implementation. This step is with respect to recognizing the common issues, necessities, or challenges around movement your AI comes to life. What common barricades may substitute your way? What activities are you able to need make beyond any doubt you provide on your AI objectives?

Change management. Finally, it's very important to think that regarding the crosscutting problems around worker impact, engagement, and communication. AI may, as an example, impact human jobs, notably if they involve automating numerous tasks or processes. What area unit the common amendment management themes that relate to your planned AI projects? At last, it's exceptionally vital you think that with respect to the crosscutting issues around laborer affect, engagement, and communication. AI may, as a case, affect human occupations, eminently in the event that they include robotizing various assignments or forms. What unit the common correction administration topics that relate to your arranged AI ventures?

2 Cross Cultural Translation

Interpretation isn't as it were an etymological act; it's conjointly a social, relate act of communication over societies. Interpretation until the end of time includes each dialect and culture basically since that cannot be isolated [4]. Dialect is socially implanted. It each communicates and shapes social reality, and so the which suggests of etymological things can exclusively be caught on once thought of on board the social setting all through that the phonetic things are utilized. Interpreters ought to pay decent consideration to varieties amid a comparable implies and degree of conventionalization interior the accessibility and target societies once exchanging a

content from one culture to a particular [5]. One of the pre-eminent characteristics of translation is its 'double-bind situation', where the interpreter got to interface the accessibility text in its social setting to the target communicative-cultural condition.

3 Contextual Artificial Intelligence

Relevant Fake Insights: The building pieces of a fruitful relationship between people and AI. The center is that the definition of a collection of needs that alters a subordinate relationship between AI and people. Talk AI needs to be comprehensibly, versatile, customizable and governable, and context-aware. Talk AI doesn't inquire a chosen algorithmic run the show or machine learning procedure—instep, it takes a human-centric studied and approach to AI [6].

The center is that the definition of a collection of needs that alter a subordinate relationship between AI and people. Talk AI should be coherently, versatile, customizable and governable, and context-aware. Here's what that shows up like inside the genuine world [7].

- Intelligibility in AI alludes to the necessity that a framework ought to be able to clarify itself, to speak to its clients what it knows, how it knows it, and what it is doing around it. Comprehensible is required for believe in AI systems.
- Adaptively alludes to the capacity of an AI framework, when prepared or conceived for a particular circumstance or environment, to be versatile sufficient, so it can be run essentially in a distinctive circumstance or environment and meet the user's desires. For illustration, a keen domestic right hand that controls my house knows my inclinations, but will it be able to interpret those to my mom's domestic when I visit?

4 Artificial Intelligence Translators

AI interpreters unit computerized apparatuses that utilize progressed computing to not solely decipher the words that unit composed or talked, but to boot to interpret the proposes that (and as a rule assumption) of the message [8]. These closes up in larger exactness and less errors than once victimization mechanical gadget interpretation. Instead of inquiring your instructor what a word or state proposes that of late you'll essentially take note Relate in nursing app that mechanically deciphers a distant off dialect to your normal dialect. Virtual collaborators like Siri utilize dialect acknowledgment code, as do learning apps like couple vernacular or Rosetta stone. This code works by victimization discourse acknowledgment calculations to spot the discourse communication [9]. It at that point investigations the sound to appear it into content. A few gigantic organizations of late have begun victimization manufactured insights in their dialect administrations indeed Confront book envelops perform for interpretation. This depends on text-to-text interpretation that changes over a word, express

or possibly a section into another dialect [10]. AI interpreters moreover can make development a full load less demanding, especially in case you're going by some place like China wherever not as it were is that the dialect completely diverse, be that as it may the letter set is as well! These apps and administrations may be a savior once talking to local people or requesting one thing to eat in an eating house [11, 12]. You'll just hold your phone over a menu to decipher the dishes and communicate with the server through your mouthpiece.

4.1 Advantages of Artificial Intelligence Translation

- Quality in domain- and language-specific engines is increased.
- Unbelievably fast, commonly only takes a moment around.
- Ideal for websites.
- AI tools can allow you to translate an outsized range of languages, whereas an individual's might only be ready to translate one or two.
- Turning into additional correct.

4.2 Disadvantages of Artificial Intelligence Translation

- Accuracy may be low in some industries.
- Mass localization is poor, with some languages giving additional correct results than others.
- AI doesn't acknowledge the context one thing is claimed in, therefore will turn out confusing results.

5 Conclusion

Interpretation can be a strategy of commutation a content in one dialect by a content in another dialect. A content isn't essentially a add up to of its components, and once words and sentences unit utilized in communication, they blend to form meaning in a few ways that. Subsequently, it's the complete text to be interpreted, rather than partitioned sentences or words. A communicative text can carry its social alternatives though moving from one language to a diverse. The interpreter got to be familiar with SL and metal societies, get it the point of the communication conjointly the gathering of people for legitimate and on-time higher cognitive handle to undertake to his/her interpretation as successful society communication. We have a propensity to go to keep intellect that, since of varieties, there's no real interpretation between any two dialects. What one will trust for is relate guess. The part of comparative the

frameworks and societies of the two dialects, the parcel of prudent the translation in society communication.

References

1. Hariyanto S (2009) The implication of culture on translation theory and practice. Retrieved to august 21
2. House J (2009) Translation. Oxford UP, New York
3. Kianush K (2009) Persian miniature. Retrieved to 30 July 2009, from http://en.wikipedia.org.
4. Ketabi S, Ordudari M (2008) Translation focus. Chaharbagh Publication, Isfahan
5. Miremadi SA (2004) Theories of translation and interpretation. SAMT, Tehran
6. Newmark P (1988) A textbook of translation. Prentice Hall, UK
7. Newmark P (1988) Approaches to translation. Prentice Hall, UK
8. Richards JC et al (1992) Dictionary of language teaching & applied linguistics. Longman, UK
9. Richards JC (2005) Interchange2 (3rded). New York: Cambridge UP. Risk Analysis for a commercial computing service. In: Proceedings of the 21st IEEE international parallel and distributed processing symposium, Long Beach, California, USA (March 2007)
10. Sulistio A, Kim K, Buyya R (2008) Managing cancellations and no-shows of reservations with overbooking to increase resource revenue. In: Proceedings of the 8th IEEE international symposium on cluster computing and the grid, Lyon, France
11. Tiwari R, Sharma M, Mehta KK (2020) IoT based parallel framework for measurement of heat distribution in metallic sheets. Solid State Technol 63(06):7294–7302. ISSN: 0038-111X
12. Shrivastava R, Singh M, Thakur R, Subrahmanya Ravi Teja KS (2021) A real-time implementation for the speech steganography using short-time Fourier transformior secured mobile communication. J Phys Conf Series 2089:012066. https://doi.org/10.1088/1742-6596/2089/1/012066

Design of Low-Power OTA for Bio-medical Applications

R. Mahesh Kumar, C. Jaya Sree, G. Rama Krishna Reddy, P. Naveen Kumar Reddy, and T. Bharat Kumar

Abstract The OTA would be a crucial element of any biological IMD. Perhaps for the first instance in the analogue chain, the biological information from the human body has been analysed, defining some critical device settings. Due to the highly restricted frequency, it also is challenging to develop a completely integrated biological data recording device. This article proposes a completely integrated OTA architecture with minimal power dissipation. It utilizes a novel mirror bulk-controlled OTA to produce low power and good gain. The OTA has quite a 45 dB connection and consumes 0.168 uW of power. The simulation has been carried out in TSMC 180 nm CMOS technology.

Keywords Biological IMD · OTA · Low power · Biological information · Analogue chain

1 Introduction

Biopotential signal monitoring and record is critical for clinical diagnosis, and modern clinical procedures necessitate regular monitoring of such information. Patients are often connected to extended recording gadgets in order to acquire bodily signals for diagnostic reasons. This impairs their movement and contributes to an overall sense of uneasiness. As a result, acquisition time is reduced and continuous patient monitoring is avoided, which has a negative effect on the overall identification of disorders.

Their goal is to create a biopower recording device that really is comfortable, offers long-term resource efficiency, has a good signal strength, and therefore can be adjusted to assist in a wide range of medical applications. Its objective is to guarantee that mini-ambulatory biopotential acquisition systems work quietly and use very little power. This is important in order to conduct a thorough investigation of the architecture of these interfaces and their eventual function.

R. Mahesh Kumar (✉) · C. Jaya Sree · G. Rama Krishna Reddy · P. Naveen Kumar Reddy · T. Bharat Kumar
Annamacharya Institute of Technology and Sciences, Rajampeta, India
e-mail: Rmahesh786369@gmail.com
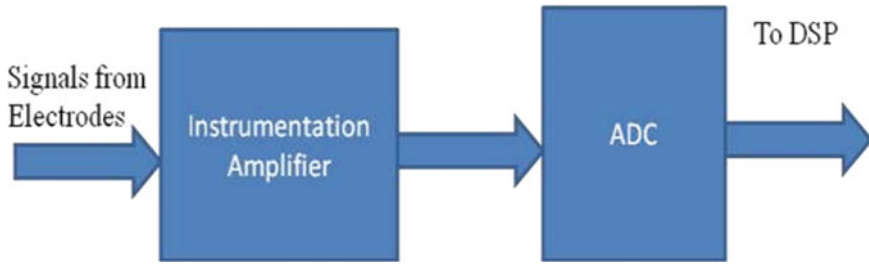
Fig. 1 Typical ECG monitoring system

Due to the unique electrical characteristics of this information, it is difficult to develop devices capable of tracking biopotential signals. Such pulses typically have an amplitude of a few mv to several mv. Once more, it will be defined by a separate frequency range ranging from sub-hertz to many hundred hertz, depending on the observed biopotential signal [1, 2].

Figure 1 illustrates the block diagram of the ECG monitoring system's analogue frontend.

It consists of an electrostatic discharge (ESD) and INAMP, an\ADC, and a wireless transmitter [2]. Defibrillator safety is typically integrated in current ECG systems prior to the analogue frontend, to safeguard the patient, the user, and the device itself during discharges or emergencies [3]. First before signal is sent to the instrumentation amplifier, it is digitized by the ADC. This is possible when a negligible total energy is dissipated [3, 4, 5, 6].

OTA consumes the most of electricity in IA [7]. This low-power OTA would be intended for use in biological applications. OTA is indeed a critical component of numerous analogue circuits and is found in a broad range of consumer devices, commercial, and scientific devices [8]. Based on the system requirements, an OTA should meet a variety of design criteria. OTA's low transconductance enables the building of low frequency filters, which are critical for capturing biological signals that are very low in frequency and amplitude necessary to construct a filter for the biopotential signal's transconductance of OTA in the nS range [9].

2   Proposed Architecture of OTA

We merged the concepts of bulk-driven transistors and designed the OTA to function in the weak inversion area, which is ideal for biological applications that need low-power operation and extended battery life. The suggested OTA is schematically shown in Fig. 2. As can be observed from the proposed OTA, input voltages drive the bulk body terminals of PMOS 1, 3 and PMOS 2, 4. Additionally, the suggested NMOS 1, 2 function as differential inputs, while NMOS 3, 6 and NMOS 4, 5 function
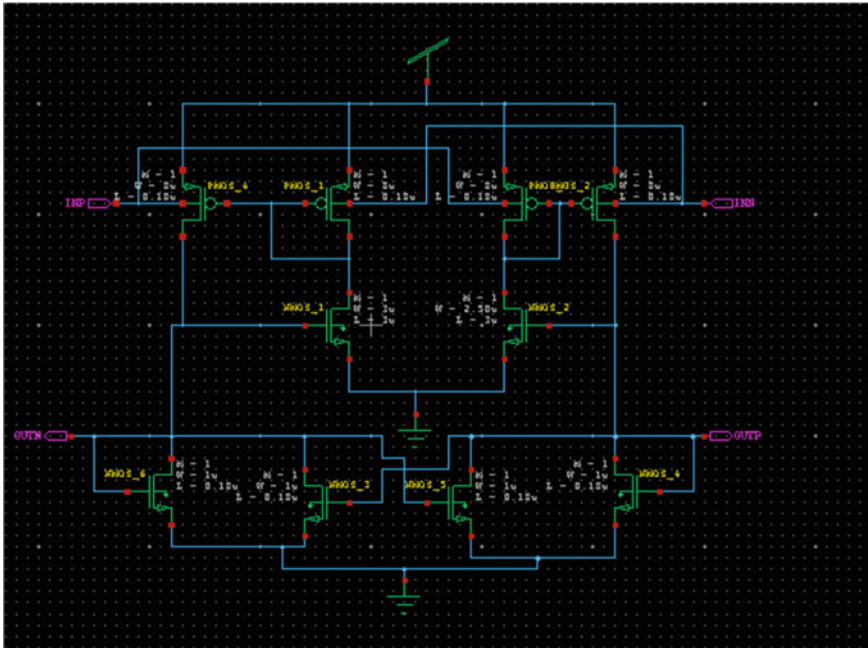
Fig. 2 Proposed architecture of OTA

as current mirrors. Initial differential amplification is performed via the differential inputs NMOS 1, 2. NMOS 3, 4, 5, 6 is used to do the final amplification.

3 Simulation Results

Tanner EDA tools are used to carry out this task on TSMC CMOS 180 nm technology. The OTA's transient analysis is depicted in Fig. 3.

The OTA's transient analysis is shown in Fig. 3. The suggested OTA is powered by a 2 m Vpp sinusoidal pulse with a frequency of 10 Hz. Additionally, the signal is amplified to a 700 m Vpp output voltage with very minimal noise.

According to the AC study, the OTA has a DC gain of 45 dB, as seen in Fig. 4. With transistors functioning or running in the weak inversion zone, power consumption is significantly decreased. Additionally, efforts are taken to limit OTA noise.

Table 1 compares the current study to previously published works as per which the recommended task consumes less energy and yields a higher gain.
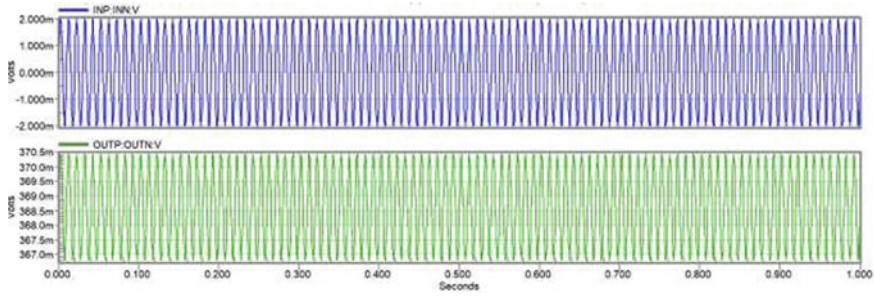
Fig. 3 Transient analysis of OTA
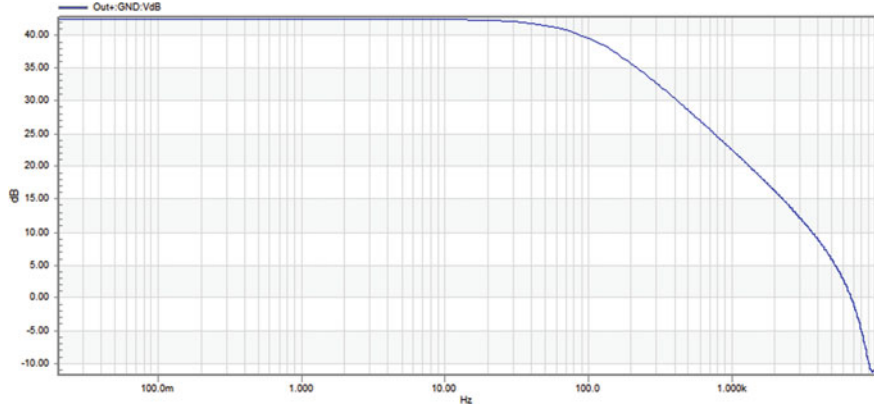
Fig. 4 AC analysis of proposed OTA

Table 1 Comparison of present work to existing work

Parameters	[10]	[11]	This work
CMOS technology used (nm)	180	90	180
Supply voltage (V)	0.9	1.1	0.8
Power consumption (uW)	1.26	7.24	0.168
Gain (db)	40	43	45

4 Conclusion

This article suggests an OTA for biomedical applications. The OTA is an indispensable part of any biological IMD. Biological data from the human body was assessed for the first time in the analogue chain, defining important device parameters. It employs a unique bulk-driven mirrored OTA to generate low power and high gain. The OTA interface is 45 dB and consumes 0.168 uW of power. The simulation is conducted using CMOS technology with a 180 nm feature size.

References

1. Haddad SAP et al (2006) The evolution of pacemakers. IEEE Eng Med Biol Mag 38–28
2. Ellenbogen KA (2006) Clinical cardiac pacing, defibrillation, and resynchronization therapy, 3rd edn. Elsevier, New York, NY, USA
3. Wong LSY et al (2004) A very low-power CMOS mixed signal IC for implantable pacemaker applications. IEEE J Solid-State Circuits 39(12):2246–2456
4. Chi YM, Cauwenberghs G (2010) Wireless non-contact EEG/ECG electrodes for body sensor networks. In: Proceedings of international conference on body sensor networks, pp 297–301
5. Guermandi M, Cardu R, Franchi E, Guerrieri R (2011) Active electrode IC combining EEG, electrical impedance tomography, continuous contact impedance measurement and power supply on a single wire. In: Proceedings of ESSCIRC, pp 335–338
6. Shaik Karimullah, D. Vishnuvardhan, Muhammad Arif, Vinit Kumar Gunjan, Fahimuddin Shaik, Kazy Noor-e-alam Siddiquee, "An Improved Harmony Search Approach for Block Placement for VLSI Design Automation", Wireless Communications and Mobile Computing, vol. 2022, Article ID 3016709, 10 pages, 2022. https://doi.org/10.1155/2022/3016709
7. Xu J, Mitra S, Matsumoto A, Patki S, Van Hoof C, Makinwa KAA, Yazicioglu RF (2014) A wearable 8-channel active-electrode EEG/ETI acquisition system for body area networks. IEEE J Solid State Circuits 49(9):2005–2016
8. Lee S-Y et al (2011) A programmable implantable microstimulator SoC with wireless telemetry: application in closed-loop endocardial stimulation for cardiac pacemaker. IEEE Trans Biomed Circuits Syst 5(6):511–522
9. Gunjan, V.K., Shaik, F., Kashyap, A. (2021). Detection and Analysis of Pulmonary TB Using Bounding Box and K-means Algorithm. In: Kumar, A., Mozar, S. (eds) ICCCE 2020. Lecture Notes in Electrical Engineering, vol 698. Springer, Singapore. https://doi.org/10.1007/978-981-15-7961-5_142
10. Kumar R, Sharma S, Singh A (2016) A low power low transconductance ota in 180 nm Cmos process forbiomedical applications. Ijcta 9(11):5427–5434
11. Tarique Ali MD, Ahmad N (2017) Design and simulation of CMOS OTA with 700mv, 60db GAIN and 10PF load. IJEDR 5(2):1359–1361

Analysis of the Efficiency of Parallel Prefix Adders

S. Fayaz Begum, M. Kavya Sree, S. Amzadhali, J. Venkata Sai Sushma, and S. Sai Kumar

Abstract ALU is at the core of each and every processing unit, as well as the adder circuit would be at the core of any ALU. The adder performs subtraction and is the fundamental component of multipliers' circuitry. Due to the fact that additions are the most basic arithmetic function and adder has become the most essential arithmetic part of the processing unit, the topic of VLSI arithmetic must have captivated numerous modern VLSI researchers over the decades. Numerous conventional and creative ways have been developed over the years to maximize the efficiency of the addition in latency, energy consumption, and area. This article reviews a few classical addition methods, evaluates their efficiency, and afterwards assesses and analyses the performance of different parallel prefix adders. These prefix adders have become currently the most effective method for DSP processors. The objective is to choose the optimal adder from among those presented for a specific application in terms of on-chip power usage, resource use, and computation speed. The adders have been designed in Verilog HDL using Vivado Design Suite on Xilinx ZYNQ-7000 series SoCs.

Keywords Adders · Zynq-7000 · Vivado design suite

1 Introduction

The calculations on VLSI chips are heavily reliant on power-efficient, high-speed, adder circuits. Traditional adders are capable of meeting this need up to a specific range of inputs. Even as input size rises, several of these adders slows down when it comes of calculation speed while using less energy as well as area, and others are quick but use a disproportionate amount of power. Parallel prefix adders provide an excellent balance of speed and power consumption [1]. This article will cover and analyse some conventional adders as well as parallel prefix adders. Additionally, it analyses their effectiveness in terms of energy consumption as well as latency,

S. Fayaz Begum (✉) · M. Kavya Sree · S. Amzadhali · J. Venkata Sai Sushma · S. Sai Kumar
Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: fayazbegums@gmail.com
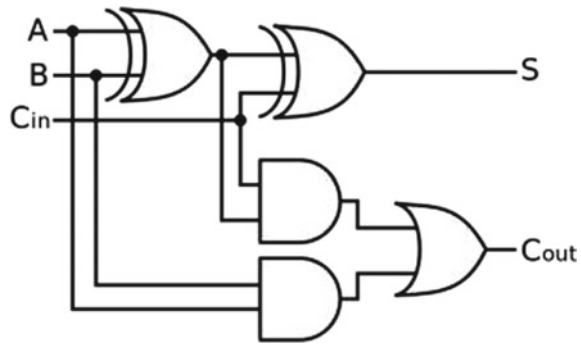
emphasizing the shortcomings of previous systems. Furthermore, the study presents and compares the power usage, area, and delay characteristics of every parallel prefix adder addressed in this research.

This section discusses the basic unit of traditional adders [2]. Section 2 explores a few common multi-bit adders, along with the RCA, the carry-bypass adder, and the CLA adder. Additionally, the part discusses the disadvantages of traditional techniques and the necessity of parallel prefix adders. Next, in Sect. 3, we will cover the principles, prefix trees, and principles underlying parallel prefix adders, specifically the Kogge-Stone, Han-Carlson, Brent-Kung, and Ladner-Fischer adders. The next part, Sect. 4, evaluates the efficiency of all the parallel prefix adders covered before, first conceptually and then empirically by using Vivado Design Suite and Verilog code generated and realized on a Xilinx Zynq.

As noted before, let us begin by discussing the basic 1-bit full-adder circuit that serves as the foundation for all conventional methods to adder circuitry.

Full Adder

A complete adder would be an extension of the half adder that can add 2 inputs and the preceding carry. As shown in Fig. 1, the circuitry has a delay of two time units. A complete adder cannot satisfy the ALU's criteria for an adder. Designers have already seen the insertion with one input up to this point, but then in real time, all inputs become multiple bits. Now, let us examine the circuitry capable of achieving addition over several bits.

2 Conventional Multi-bit Adders

Time delay is perhaps the most critical efficiency characteristic for all circuit design. Arithmetic circuitry now becomes component of the circuit's datapath. The datapath must be as quick as feasible, with energy consumption becoming a secondary consideration. The critical path is indeed the delay between any two input and output entries
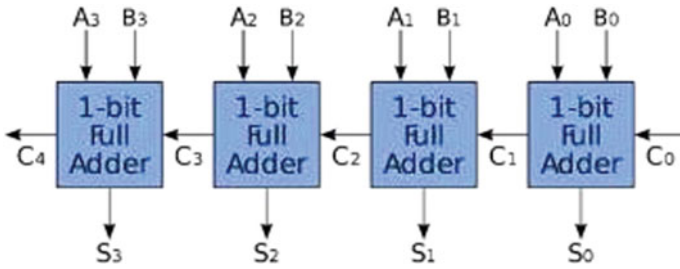
Fig. 2. 4-Bit ripple-carry adder

that is as lengthy as feasible. This is crucial because adders apart from parallel prefix adders may readily be analysed and assessed using the critical route technique.

2.1 Ripple-Carry Adder

This is a binary adder which can compute the mathematical addition of two binary integers. As implied by the circuit's name, a ripple would be created for the carry. This is formed by cascading 'n' full adders with n-bit adders, while each full adder's output carry is linked to an input carry of a following full adder (Fig. 2). The circuit has a delay of 'n' time constants for just 'n' bit adder, in which the latency of a single complete adder is assumed to be one constant value. Here, the critical path begins with the first bit and continues to the final bit, because each full adder must wait for the preceding full adder to carry before performing individual bit addition.

This adder is not commonly recommended due to the significant delay it introduces as the number of data bits grows.

A truth table for just a complete adder is seen in Table 1, but from a different viewpoint. This novel strategy is the very first step in developing more efficient adders. Sum and Cout are defined here in relation of G (generate), P (propagate), and D (delete). Those three pieces characterize the carry-out and therefore are introduced since they are inextricably linked to Cinistight. G, P, and D are independently exclusive variables and thus are defined:

Cout is always 1, no matter what Cin is. $G = A\&B$ Cin gets cloned into Cout (propagated). $P = A^\wedge B$ Delete: Cout is always equal to zero, regardless matter what Cin is. $D = (\sim A)\&(\sim B)$ Cout may be represented as follows: Cout $= G/(P\&\text{Cin})$.

2.2 Carry-Skip Adder

Figure 3 illustrates a five-bit RCA [3] with create and propagated logic at each bit. Additionally, the chain's ultimate execution is multiplexed across two cases. When

Table 1 Truth table of a full adder with generate and propagate bits

A	B	Cin	Cout	S	D	G	P
0	0	0	0	0	0	1	0
0	0	1	0	1	0	1	0
0	1	0	0	1	0	0	1
0	1	1	1	0	0	0	1
1	0	0	0	1	0	0	1
1	0	1	1	0	0	0	1
1	1	0	1	0	1	0	0
1	1	1	1	1	1	0	0
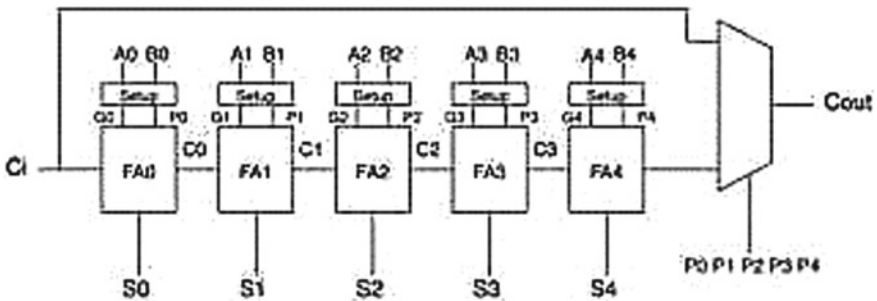
Fig. 3 Five-bit RCA using *G* and *P* bits with a 2 × 1 Mux [3]

all 5 positional propagates become true, the 5-bit adder as a whole will propagate. The carry-out becomes picked as the carry-in in this scenario. Otherwise, the carry-out is the result of the 5-bit RCA. This strategy does not seem to significantly reduce crucial route latency. To begin, the final output of the 5-bit adder becomes S4 instead of Cout. Therefore, even though the bypass adder succeeds in producing Cout sooner, the sum bit must still be waited for. Furthermore, the condition in which all bits propagate is a particular case. *P*0, *P*1, *P*2, *P*3, and *P*4 have 32 possible values, and '11111' is simply one of them. Therefore, although if Cout were to be regarded our final output, we would see an improvement in its latency just 1 out of every 32 occasions. While the finest delay has been improved, the very worst delay remains 4Tcarry + Tsum.

2.3 Carry Lookahead Adder

As demonstrated previously, propagate logic and generate logic are utilized to create quicker adders. The CLA [2] takes use of these to compute one or several carry bits prior to the addition. This decreases the time required to compute the output of the adder's bigger bits. After generating the carry with all bits, the total bit is determined

using the XOR operation of propagate as well as carry. Thus, this delay maintains consistent including all *n*-bit adders, corresponding to three-time constants; the three levels denote propagate as well as generate, carry-lookahead generator blocks, and then final stage sum. The primary downside would be that the adder consumes a lot of space, and as the amount of data rises, the area used climbs exponential, requiring more capacity to operate the circuit.

To summary, standard adder circuit techniques do not provide optimal trade-offs between latency, energy, and area. Although the RCA needs no additional hardware, it is indeed essentially sluggish due to the possibility of propagating the intake carry all ways to the outputs. As a result, the delay grows linearly with bit size of the input. But at the other hand, the CLA executes addition very quickly, and the latency does not really scale with the input width. Nevertheless, hardware consumption steadily rises, as does the energy required during addition. Surprisingly, the CSA outperforms the RCA in terms of speed, only when the optimum K and N are used. Apart from that, this adder has a comparable latency to the RCA that might not be consistent with the process models used to manufacture the VLSI chips, since not all bit configurations are best for reducing latency when employing a CSA. For instance, we could boost speed only by cascading five 5-bit RCA to build a 20-bit adder, like stated earlier. In the next part, we will explain the principles of parallel prefix adders, that provide favourable trade-offs among latency, power, and area.

3 Parallel Prefix Adders

In the preceding sections, we examined classic adders and determined that the CLA was the quickest while using a disproportionately high amount of logical and power. Nevertheless, the biggest feature of CLA is that when conducted hierarchically, it enables the creation of a few of the quickest adders imaginable while using less logic and power than the typical CLA.

Considering the 3rd bit's carry expressions,

$$C2 = G2|P2\&G1|P1\&P2\&G0|P0\&P1\&P2\&\text{Cin}$$

We may see a pattern within those expressions, if the first word in the preceding carry expression indicates the actual position, so producing a carry. The second term denotes the preceding location that generated the carry as well as the subsequent position that propagated it. This third term evaluates the probability of a carry-into the first bit position, which propagates all following positions; the final period examines the chance of a carry-into the first bit position, which propagates all following positions. The final term is referred to as the 'group-propagate' phrase because it refers to the entire bit locations that propagate the input carry to an output. All the other words in these formulations are collectively referred to as 'group-generate' terms because they reflect the production of a carry at each appropriate bit position and afterwards transmitted to the output. Group propagates and generates may be specified for any

number of bit locations, for example 2–4, but they must be contiguous. *Pxy* and *Gxy* are used in this study to denote group propagate and generate for the *x:y* range [4].

The principle of group generate and propagate, in conjunction with the 'dot' operator, enables the building of very effective adders [5] dubbed 'parallel prefix adders'. The dot operator should be used in parallel adder architecture to create a longer group of two contiguous sets of generates and propagates. As such, it is described as.

$$(Gxz, Pxz) = (Gxy, pxy)o[G(y + 1)z, P(y + 1)Z]$$
$$= [G(y + 1)z| Gxy\&P(y + 1)z\&Pxy]$$

The dot operator is really not symmetric; this should join uninterrupted ranges while disregarding any redundancies in the input range.

Now that we have defined CLA, group produces, and propagates; we may create and evaluate a new category of adders dubbed parallel prefix adders [6]. Even as name implies, such adders accomplish their speed through with a combination of parallelism as well as a hierarchy approach to calculating the carry of every bit position. They do, however, take up more space than a RCA and have the drawbacks of overloading on gate outputs (fan-out) and complicated routing. There have also been five significant parallel prefix adders, within each set of benefits and limitations. These are the Kogge-Stone multiplier, the Brent-Kung multiplier, the Han-Carlson multiplier, the Ladner-Fischer multiplier, and the Knowles multiplier. Notice that the architecture of Knowles adders also is not explored in this work because it would need the inclusion of classes and therefore would lengthen the text. The next section discusses the Kogge-Stone adder, which may be the smallest of the parallel prefix adders.

3.1 Kogge-Stone Adder

It is a widely used parallel prefix adder. It does have a quicker design that increases computing speed; however, the circuit size is larger. It generates the adding in three phases [7], like every parallel prefix adder does: pre-processing, carry graph generation, and then post-processing (sum block). The carry graph network is shown below in Fig. 4.

Just at 0th stage, generate (*gi*) as well as propagate (*pi*) functions are constructed for every bit based on the inputs *A/B*. Carry becomes derived just at following step, which would be the carry generate phase, using group propagate as well as generate. The number of stages in this stage changes in proportion to the amount of input bits. The last step generates the total; here, the propagate (*pi*) is being XORed with carry created by the carry system (*ci*). The adder's merits are that it does have a low fan-out, a rapid addition process, and a minimal circuit delay, which would be determined by

Fig. 4 Prefix tree of 8-Bit
KS Adders [3]
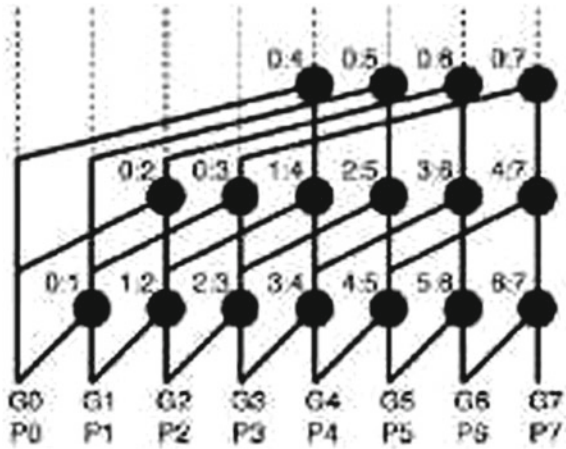
$\log 2(n)$, wherein n represents the size of bits in the inputs. The drawback is that the adder consumes a lot of space.

3.2 Brent-Kung Adder

As demonstrated in the carry graphs network, one of the downsides of the KS adder would be the increased number of wires. This is a large number that expands exponentially in proportion to the magnitude of the inputs. Additionally, the cables must cross several other lines, complicating routing. Consider the 8-bit KS adder with the assumption that only the last carry is significant. In this example, the majority of dot operations may be eliminated, resulting in the following tree diagram of Fig. 5.

This reduces routing congestion, and it is not a true 8-bit adder, because all internal carries remain required to compute S0 through S6. As a result, we modify this trimmed tree to create the Brent-Kung adder [8]. Four more dot operators just at top of last carry are required for an 8-bit BK adder. This is used to compute the necessary carries with the fewest extra hardware units feasible. Figure 6 is a prefix tree diagram for an 8-bit BK adder.

As seen, BK adders are significantly smaller than KS adders. The 8-bit BK adder contains 11 dot operations, while the KS adder contains 17. The 16-bit BK adder contains 26 dot operators, while the 16-bit KS adder contains 49. Thus, the area represents the BK adder's advantage, which becomes increasingly apparent as the input size increases. In contrast to the KS adder, this BK adder's wiring remains likewise readily routable. Nevertheless, the BK adder contains two significant drawbacks. To begin, the number of steps required to compute all internal bit carries has indeed been raised. Secondly, most dot businesses have expanded beyond two units.
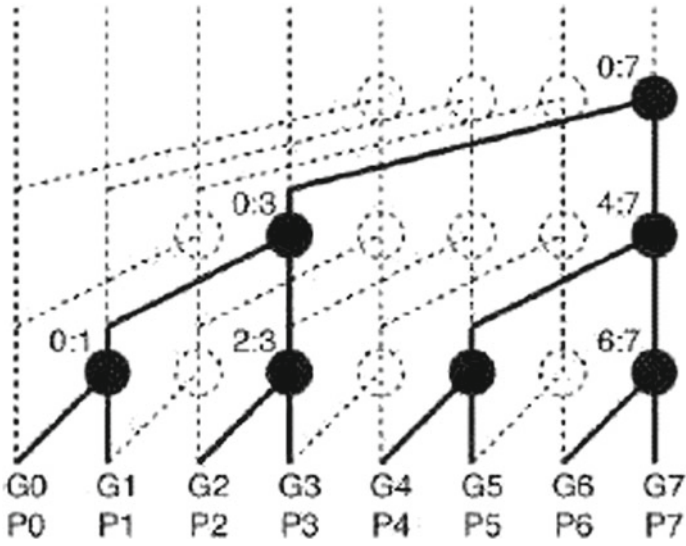
Fig. 5 Pruned tree diagram for 8-bit KS adder [3]
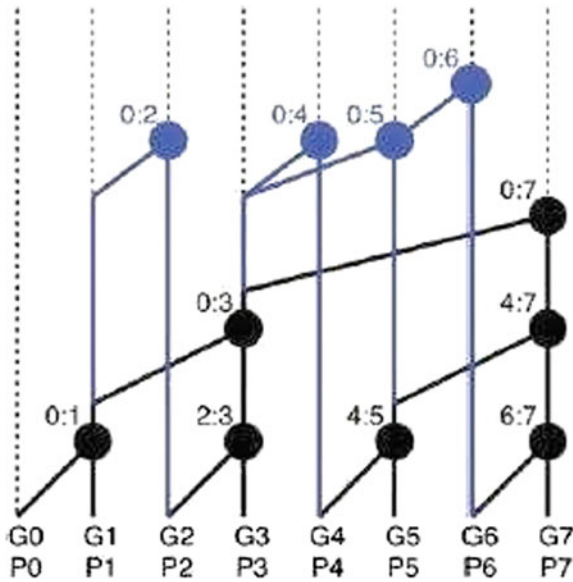
Fig. 6 Prefix tree for 8-bit BK adder [3]
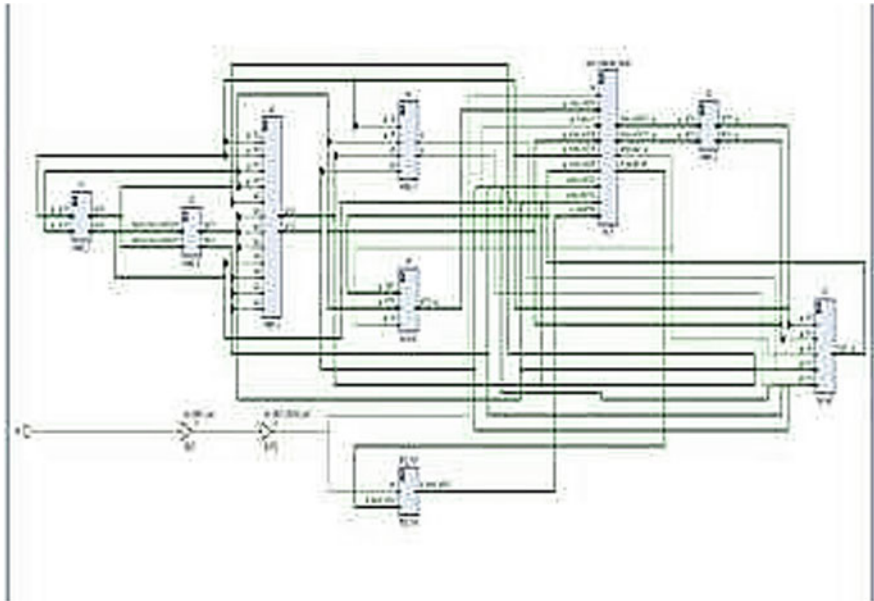
Fig. 7 Synthesized schematic of 8-bit Han-Carlson adder

3.3 Han-Carlson Adder

The Han-Carlson adder would be a hybrid technique [9] that combines the Kogge-Stone and Brent-Kung adders. It employs a BK method (slow but tiny) during the initial and final phases of carry production and a KS approach (quick but huge) during the intermediate stages. Han-Carlson's prefix tree concept is similar to Kogge-Stone's prefix tree concept in that it has a maximum fan- out of two. The distinction is that the Han-Carlson prefix tree consumes far less hardware and wire than the Kogge-Stone adder. One more logic level is required. As a consequence, this adder provides an excellent trade-off across fan-out with logic unit count.

In the above picture, the Han-Carlson adder employs a single Brent-Kung level at the start and the end of the graph, as well as the Kogge-Stone technique in the centre. Simply said, the Han-Carlson adder accomplishes faster speeds while using less power and taking up less space. Figure 7 depicts the Vivado Design Suite-generated schematic for an 8-bit Han-Carlson adder.

3.4 Ladner-Fischer Adder

Other adder seems to be the Ladner-Fischer Adder [10], which is also the quickest adder in terms of design effort as well as the most often used high-performing adder
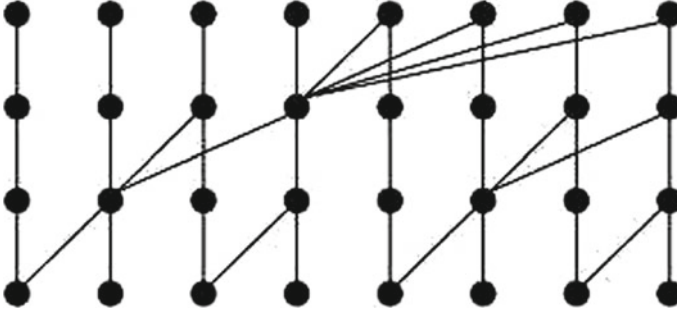
Fig. 8 Prefix tree for 8-bit Ladner-Fischer adder

with Kogge-Stone adder. As with every parallel prefix adder, it generates the adding in three stages: pre-processing, carry graph structure, and post-processing.

Just at 0th stage, generate (gi) and propagate (pi) functions are constructed for every bit based on the inputs A and B. Therefore, the following step, carry generate, has been derived using group propagate plus generate carry. The number of cycles in this stage changes in proportion to the amount of input bits. The carry network carry graph in Fig. 8 illustrates the principles used to generate carry.

The last step is where the total is generated. In this case, propagate (pi) seems to be XORed with carry created by the carry network (ci). The benefits here are that the logic level is generally basic, which speeds up the circuit and reduces its size. The downside is that, in comparison to certain other parallel prefix adders, it does have a huge fan-out. Figure 9 depicts the Vivado Design Suite-generated schematic for just an 8-bit Ladner-Fischer adder.

4 Performance Comparison of Parallel Prefix Adders

After comprehending numerous parallel prefix adders with their associated prefix diagrams, it is indeed time to evaluate them [11] to determine which one is the greatest fit for a given application. Table 2 summarizes the theoretical performance characteristics of all the PPAs considered in this study as a function of the input bit size. The table additionally contains a subclass of Knowles parallel adders [12] for the purpose of aggregating the evaluation.
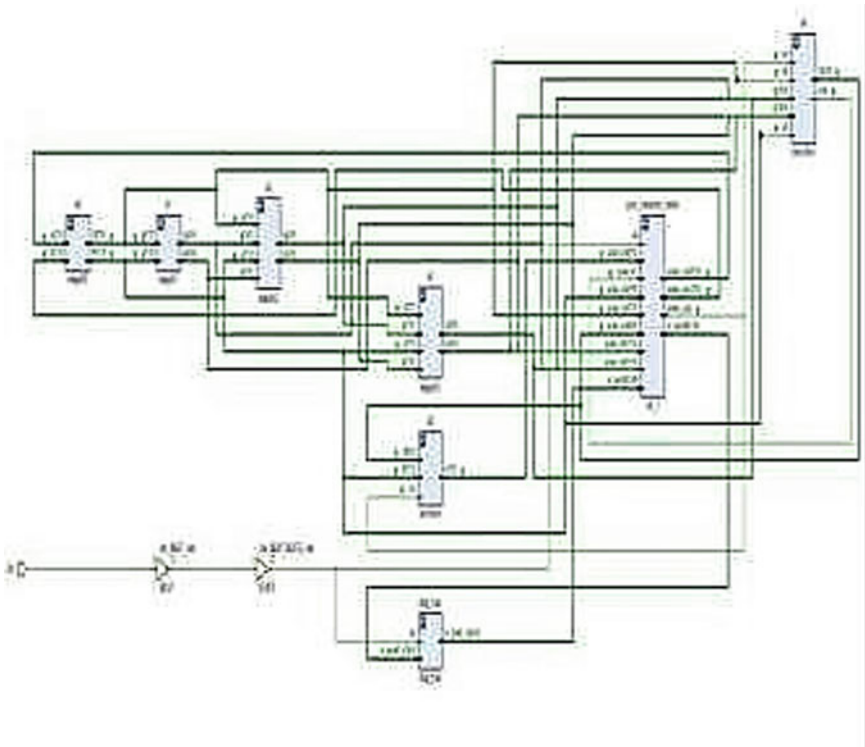
Fig. 9 Synthesized schematic of 8-bit Ladner-Fischer adder

Table 2 Theoretical comparison of PPAs

Prefix structure	Logic levels	Prefix count	Fan-out
Kogge-Stone	Log n	$N \log n - n + 1$	2
Brent-Kung	$2 \log n - 1$	$2(n - 1) - \log n$	2
Lander-Fischer	Log n	$(n/2)\log n$	$(n/2) + 1$
Han-Carlson	Log $n + 1$	$(n/2)\log n$	2
Knowles [2,1,1,1]	Log n	$N \log - n + 1$	3

4.1 Analysis of Efficiency Utilizing Xilinx's Vivado Design Suite

Xilinx offers FPGAs and other development boards as addition to hardware developer tools such as the Vivado Design Suite [13]. The suite addresses the necessity for design flows and offers a tried-and-true technique for maximizing efficiency from idea through implementation and debugging. The design flow and technique for implementing and comparing the designs are shown in Fig. 10.
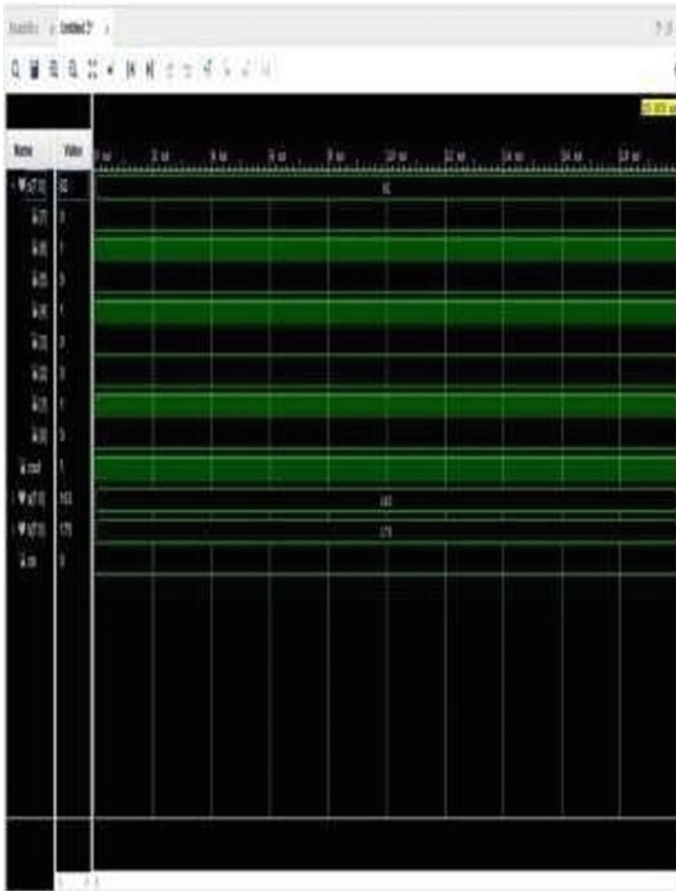
Fig. 10 Simulation results for 8-bit Kogge-Stone adder

The designs were implemented on the Zynq Evaluation and Development Kit (ZedBoard) [14] that includes an APSoC called the XC-7Z020 CLG484 APSoC (Zynq-7000 series). It is equipped with a dual Cortex A9 Processing System (PS) and 85,000 Programmable Logic Cells (PLCs) (PL or FPGA).

Table 3 covers all of the previously described performance characteristics, as well as delay and static power usage. Additionally, it demonstrates that the Kogge-Stone adder has significantly shorter latency than other prefix adders also as bit width of input grows.

Table 3 Summary the performance comparison of prefix adders

Name of the adder	Static power watts	Dynamic logic power watts	Delay (ns)	No. of logic LUTs
Kogge-Stone	0.165	0.911	26.895	37
Han-Carlson	0.158	0.843	27.091	32
Brent-Kung	0.150	0.747	27.317	20
Ladner-Fischer	0.156	0.838	26.902	28

5 Conclusion

We have explored the drawbacks of traditional adders and the ideas underlying parallel prefix adders throughout this work. The development of parallel prefix adders utilizing Vivado Design Suite solidified our grasp of their trade-offs. According to the acquired data, the Kogge-Stone adder is the quickest adder despite using more on-chip power and space than the others. This is an excellent companion to a multiplication in a MAC unit for signal applications that require greater speed operation, whereas the Brent-Kung adder is the optimal solution for low-power VLSI designs and is less resource-intensive. The Han-Carlson adder demonstrates to become a hybrid technique, with delay, area, and power values that are intermediate among those of the Kogge-Stone and Brent-Kung adders. Furthermore, although the Ladner-Fischer adder surpasses the Han-Carlson adder in terms of resource and power usage, it is criticized for having a large fan-out. They want to expand the bit width to 32/64 bits in the future and compare results at various operation temperatures. Additionally, we would want to improve the adder circuits' dynamic and static power consumption utilizing low-power VLSI design methodologies.

References

1. Oklobdzija VG, VLSI arithmetic. University of California. [Online] Available: http://www.ece.ucdavis.edu/acsel
2. Parhami B, Computer arithmetic. Oxford University Press. ISBN: 978-0-19-532848-6
3. Abbas K, Handbook of digital CMOS technology, circuits and systems. Springer, ISBN: 978-3-030-37194-4
4. Kanaka Durga T, Nagaraju CH (2015) Implementation of carry select adder with Area-Delay-Power and efficiency. Int J Sci Eng Tech Res 11916–11920
5. Veeramachaneni S (2015) Design of efficient VLSI arithmetic circuits. International Institute of Information Technology, Hyderabad. [Online] Available: https://shodhganga.inflibnet.ac.in/handle/10603/45300
6. Kilburn T, Edwards DBG, Aspinall D (1959) Parallel addition in digital computers: a new fast carry circuit. In: Proceedings of IEE, vol 106, pt. B, p 464
7. Kogge P, Stone H (1973) A parallel algorithm for the efficient solution of a general class of recurrence relations. IEEE Trans Comput C-22(8):786–793
8. Brent RP, Kung HT (1982) A regular layout for parallel adders. IEEE Trans Comput C-31(3):260–264

9. Han T, Carlson D (1987) Fast area-efficient VLSI adders. In: Proceedings of 8th symposium computer arithmetic, pp 49–56
10. Ladner R, Fischer M (1980) Parallel prefix computation. J ACM 27(4):831–838
11. Choi Y (2004) Prallel prefix adder design. The University of Texas at Austin. [Online] Available: https://repositories.lib.utexas.edu/handle/2152/1300
12. Knowles S (2001) A family of adders. In: Proceedings of 15th IEEE symposium computer arithmetic, pp 277–281
13. Xilinx (2020) Vivado design suite user guide, UG904 (v2020.1). [Online] Available: https://www.xilinx.com/support/documentation/sw_manuals/xilinx2020_1/ug904-vivado-implementation.pdf
14. AVNET (2014) ZedBoard hardware user's guide, v2.2. [Online] Available: http://zedboard.org/sites/default/files/documentations/ZedBoard_HW_UG_v2_2.pdf

Design of Efficient 8-Bit Fixed Tree Adder

**K. Riyazuddin, D. Bhavana, C. Chamundeswari, M. S. Firoz,
and A. Leela Sainath Goud**

Abstract   Adder tree would be the parallel arrangement of adders. The typical (AT) adder tree has both full-width and fixed-size adder trees. The present technique truncates the input data to facilitate the building of the fixed-size tree and to provide better precise output. RCA has been utilized in the construction of the adder tree. The suggested design is now more efficient than previous designs, as well as estimates the result nearly as accurately as post-truncated fixed-width AT. This work is primarily concerned with the replacement of adders in order to minimize latency. As the RCA tree is typically presented, in this article, the RCA is replaced with the CLA in order to reduce the critical route time required to attach the hardware of the adder tree and also to recommend an optimized framework of the adder tree using the CLA.

1   Introduction

Mobile devices with DSP technology need low-power, small-area designs to provide optimal results. DSP methods must be incorporated in fixed-point VLSI design. Such designs need the usage of an adder tree (AT); this is a kind of tree that is often employed in parallel architectures. The form of an adder tree would be distinct from that of a SAT. As a result, the word length increases in a different pattern on the SAT as well as AT. Multiplier works on incomplete goods. As just a result, we favor adder trees to simplify complicated design. However, when comparing schemes, the constant size adder trees as well as multiplier appear inappropriate due to the fact that perhaps the fixed-size AT creates a modified version of matrices than the fixed-size multiplier. Nevertheless, the FX-AT and FL-AT have been routinely truncated using direct truncate and post-truncated procedures. Direct truncation (DT) truncates one lower order but of the each adder outputs of a full-width adder tree, while post-truncation truncates the final adder result of FL-lower AT's adder bits. Nevertheless, to get correct results, such adder trees are built using RCA. It really is required to

K. Riyazuddin (✉) · D. Bhavana · C. Chamundeswari · M. S. Firoz · A. Leela Sainath Goud
Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: krz@aitsrajampet.ac.in; shaik.riyazuddin7@gmail.com

take a novel approach to creating effective FX-AT designs, since this technique is presently lacking in the literature. A well-designed FX-AT circuit will undoubtedly affect the effectiveness of specialized VLSI systems employing sophisticated DSP algorithms [1, 2].

Multiplexers and adders are included in the carry choose adder. Furthermore, providing input signal and ground to the two-bit adders saves time. And for all multiplexers, the selection line is the single resultant carry that determines other carry input. And this action of the carry chooses adder results in a shorter hardware delay. As a result, the RCA inside the adder trees is replaced with a carry select adder to achieve the decreased latency. As this is the central argument of the proposed study. However, the area inhabited is more than that of the RCA. This space is saved by reducing device measurement. Additionally, there is a temporal delay when RCA is compared to CLA inside an adder tree. This indicates that the delay was raised from adder onto adder in order to achieve a fair delay shrink and maintain the same accuracy of output. As a result, we recommend carrying a selected adder to get access to the adder tree. While adders such as carry look and others have a smaller latency, they take up more space in the hardware architecture than CLA does. Thus, whenever evaluated to our suggested adder trees, CLA has a shorter latency and a smaller assent area than RCA. Fixed-width (FX-AT) and full-width (FL-AT) designs are trimmed to increase the output speed [3]. Nevertheless, we are compensating the shortened adder trees with compensatory input. As a consequence, a biased estimated formula would be necessary to compensate for the adder's disregarded input, which results in estimated adder adding with in design. Direct truncate and post-truncate processes do not result in efficient adder tree architecture. Thus, approximation adders are often used to replace less important bits [4].

The paper's key focus is the development of suggested adder trees via substituting CLA for RCA, thus reducing the delay and suggesting a more effective adder tree in order to get an outcome with such a smaller delay.

2 Literature Survey

Portable electronics devices progressively include minimal power, space-efficient, as well as high-performance computer platforms. DSP techniques are implemented with fixed-point VLSI devices for these kinds of applications. Adder tree (AT) is a data structure that is frequently used in parallel processing for inner product calculation and matrix–vector multiplication. Additionally, the multiplier architecture incorporates a shift adder tree (SAT) for the purpose of accumulating partial product values. Word-length increase is a frequent occurrence while doing multiplication and additions in fixed-point computing. SAT's bit matrix would be shaped differently from ATs. As a result, the word length increases in a different sequence on the SAT as well as AT. Additionally, a few additional bits would be included in the SAT to account for negative multiplier partial products. Specific solutions have indeed been offered for implementing fixed-width multipliers efficiently and with a minimum of truncation

error. Unfortunately, the approach utilized in fixed-width multipliers is incompatible with developing a fixed-width AT design owing to the fact that bit-matrices have various shapes [4].

For each N-point input signal, a full-width AT design generates a $(w + p)$-bit output sequence, wherein $p = \log2 N$. The FX-AT architecture generates w-bit output for same size input vector. Traditionally, the FX-AT design would be derived directly from of the FL-AT architecture by immediate or posttruncation. Direct truncation (DT) truncates one lower order bit of the each FL-AT adder result, while post-truncation truncates p lower order values of FL-AT's last adder outcome. Various systems for approximation addition employing a ripple carry adder (RCA) have been proposed recently to conserve power and delay. On the basis of approximation logic, a bio-inspired OR adder has been presented. The authors offer four distinct kinds of approximation adder designs [5].

Overall, adder is presented for computing the trimultiplicand without requiring carry propagation. Including some loss of precision, these approximation designs may be utilized to execute RCA with less latency and size. Using post-truncation, the approximation RCA design may be utilized to achieve fixed-width AT. Nevertheless, the approximated APX-FX-AT is inefficient in terms of area, delay [6, 7].

For strategies and improve of shifting operations, a bit-level optimisation of FL-AT with multi-constant multiplying (MCM) is presented. We present an effective FL-AT architecture based on the approximation adder for inaccurate Gaussian filter implementation in image processing tasks. We discover that the optimized AT is MCM-based design-specific, and none of the available designs addresses the challenges associated with fixed-width deployment of AT. It is noted that neither direct truncation nor posttruncation approaches efficiently construct FXATs. A novel strategy to generating efficient FX-AT design is required, that is presently lacking in the literature [8, 9].

3   Architecture of FL-AT and FX-AT

A technique for efficiently developing an FX-AT design using shortened input as shown in Fig. 1. Use of shortened inputs in FX-AT has two benefits: (1) it reduces the area and latency inside the FX-AT by reducing the adder size, while (2) it provides an opportunity to improve other processing blocks that exist ahead of the AT inside a complicated architecture. Nevertheless, the reduced input generates a significant amount of inaccuracy into the FX-AT output that must be correctly biased [5, 6, 10].

Fig. 1 Fixed-width adder for input equals 8
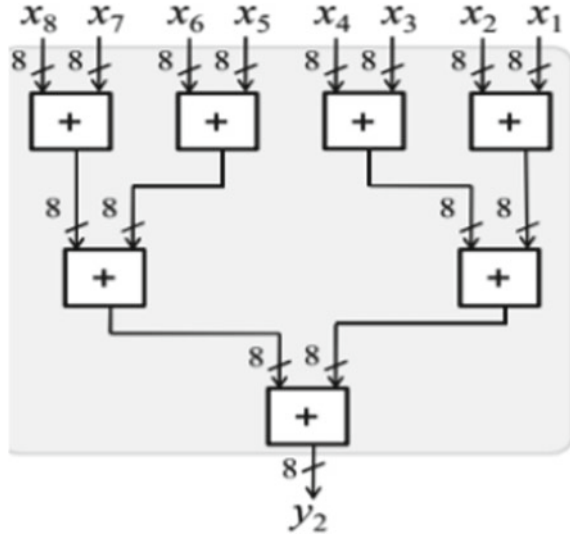
4 Architecture of Truncated Fixed-Width Adder Tree (TFX-AT)

Both the full size and fixed-size adder trees have indeed been shortened to provide an efficient design with minimal latency. To begin truncation, one first must compute the value of columns that should be abbreviated in the matrix. To do this, we will use a stochastic technique to simulate the fixed bias as shown in Fig. 2. We will remove the final three columns that will be utilized as the LSP, by utilizing formulas. Again, with the recommended truncated fixed-size adder tree, we will compute the MSP and then use the supplied equations to add the estimated input to the adder tree. TFX-AT produces the following output [5, 11, 12].

Since 5 bits have been used as input with in preceding illustration and LSP, $N = 3$ would be used for error compensating.

5 Improved ITFX-AT Architecture

In ITFX-AT, we have included half adders as well as a customized half adder that allows us to compute the approximated formula more effectively. It does, nevertheless, provide an estimate method for MSP. As contrasted to certain other adder trees, ITFX-AT is perhaps the most effective AT. These are constructed in accordance with RCA specifications. However, to increase the rapid reaction, RCA usually substituted by CLA. Whenever CLA is combined with another half adder or customized adder, its delay increases. A half adder is generally denoted by A, whereas a customized
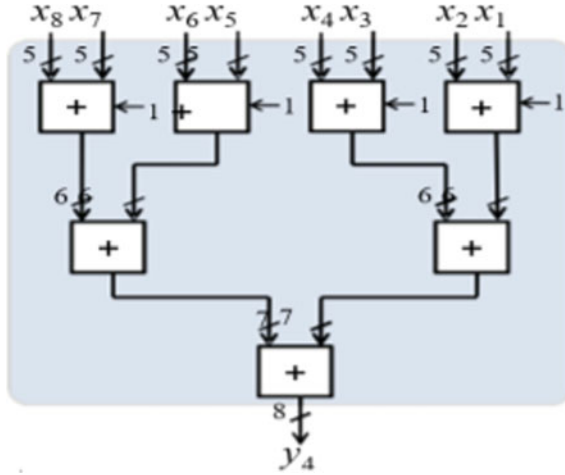
Fig. 2 Design of truncated fixed tree adder

half adder would be denoted by A*. Additionally, the adder tree is depicted below in Fig. 3.

Carry bits provided to the MSP with ITFX-AT have been taken from the LSP as shown in Fig. 4. LSP has been subdivided into MJP (major portion) and MNP (minor part) throughout this section. Furthermore, MJP is included into MSP.

Because MSP requires bias, MJP needs only a half adder as well as a modified adder to get an approximate adder. MJPs would be sent via the needed half adder and then further customized half adder to obtain the desirable result.
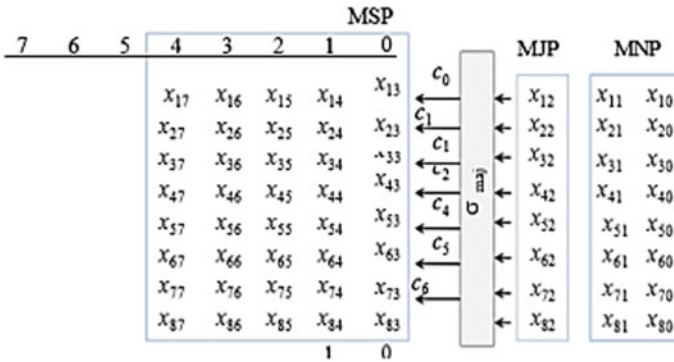
Fig. 3 Matrix for ITFX-AT
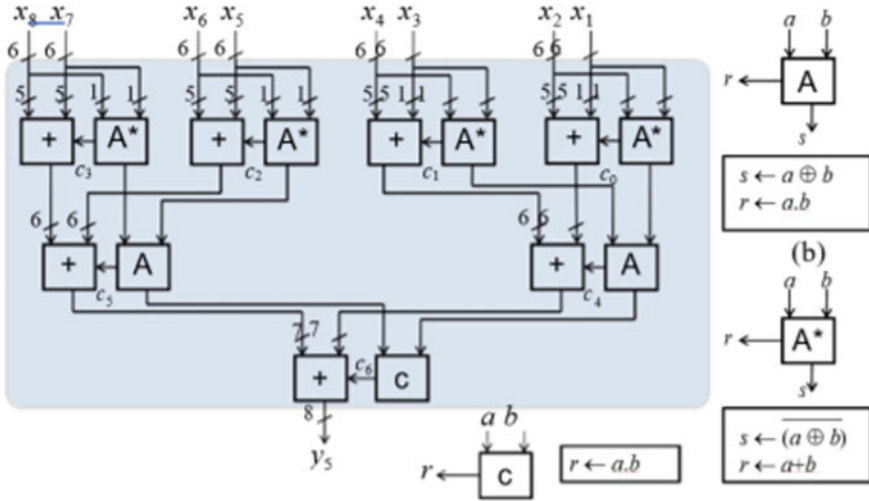
Fig. 4 ITFX-AT design utilizing half adders and customized half adders

6 Simulation Results

ITFX-AT is simple to create and obtains output very readily than truncate adder tree. Nevertheless, we are improving the FX-AT by including a half adder as well as a customized adder tree. Whenever CLA has been employed, these adders slow the rate at which output gets generated. As a result, a TFX-AT tree was used to achieve minimum latency. The result obtained for 8 bits indicates that the delay increases as the number of input bits increases. There is a significant variation in critical path latency comparing CLA vsRCA as that of the quantity of bits increases. And in terms of effectiveness, TFX-AT would be superior to ITFX-AT in terms of production. Thus, TFX-AT outperforms ITFX-AT in terms of critical path latency. The simulations were done using the Vivado design suite, as seen in Fig. 5. From Table 1, it is evident that proposed architecture exhibits better results.

7 Conclusion

In this work, RCA is substituted with CLA in AT to achieve a reduced delay. The adder tree is primarily concerned with the area and time required to get the result. As a result, we examined several forms of truncated adder trees. Truncated adder tree depends on the suggested techniques. According to the suggested system, the shortest path to the output is through a truncated fixed-width adder tree.
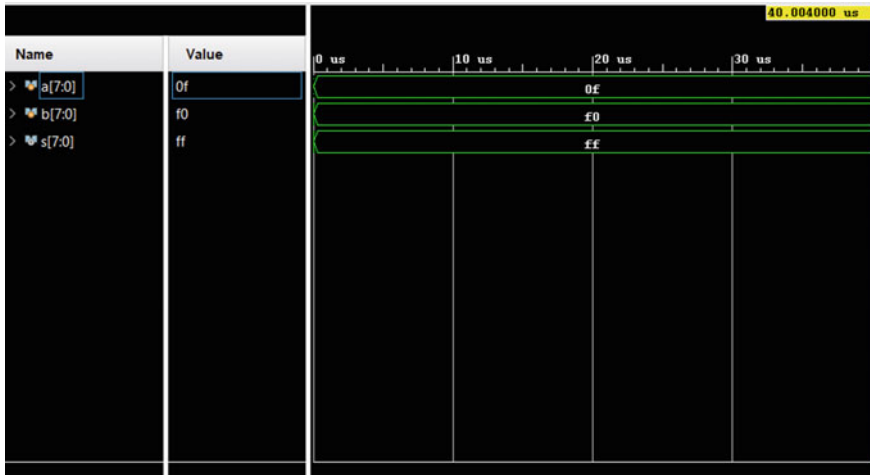
Fig. 5 Simulation result of 8-bit fixed adder

Table 1 Comparison table for 8-bit word length

Design type	Delay (ns)	Power (uW)
FX-AT-PT	0.88	379
FX-AT-DT	0.85	320
Proposed	0.74	294

References

1. Mohanty BK, Tiwari V (2014) Modified probabilistic estimation bias formulation for hardware efficient fixed-width Booth multiplier. Circ Syst Signal Process 33(12):3981–3994
2. Mahdiani HR, Ahmadi A, Fakhraie SM, Lucas C (2010) Bioinspired imprecise computational blocks for efficient VLSI implementation of soft-computing applications. IEEE Trans Circ Syst-I Regul Pap 57(4):850–862
3. Gupta V, Mahapatra D, Raghunathan A, Roy K (2013) Low-power digital signal processing using approximate adders. IEEE Trans Comput Aided Design Integr Circ Syst 32(1):124–137
4. Liang J, Han J, Lombardi F (2013) New metrics for the reliability of approximation and probabilistic adders. IEEE Trans Comput 62(9):1760–1771
5. Jiang H, Han J, Qiao F, Lambardi F (2016) Approximate radix-8 Booth multipliers for low-power and high-performance operations. IEEE Trans Comput 65(8):2638–2644
6. Pan Y, Meher PK (2014) Bit-level optimization of adder-trees for multiplie constant multiplications for effcient FIR filter implementation. IEEE Trans Circ Syst-I Regul Pap 61(2):455–462
7. https://healthtipsfr.blogspot.com/2017/04/blog-post-40.html
8. Julio RO, Soares LB, Costa EAC, Bampi S (2015) Energy-efficient Gaussian filter for image processing using approximate adder. In: Proceedings of international conference on electronics, circuits and systems, pp 450–453
9. http://www.cosasexclusivas.com/2014/06/daily-overview-el-planetatierra-visto.html

10. https://homepages.cae.wisc.edu/ ece533/images/
11. http://seamless-pixels.blogspot.in/2014/07/grass-2-turf-lawn-greenground-field.html
12. https://www.decoist.com/2013-02-28/flower-beds/

Design of T Flip-Flop Based on QCA Technology

Sudhakiran Gunda, Lakshmi Priya Mukkamalla, Venkateswarlu Epuri, Pavani Pureti, Rama Subba Srinivasulu Bathina, and Sunilkumar Bodduboina

Abstract Computational experts have been intrigued to QCA nanotechnology owing to its notable characteristics including such low energy consumption and tiny size. Different reports have been conducted inside the study outlining the usage of this approach for optimizing the process parameters of numerous QCA circuits and logic gates. The T flip-flop, a critical component of digital design, would be used to create counters and registers. This article describes an optimum implementation of an unique T flip-flop structure. To validate the developed circuits and provide the simulation results, the QCADesigner program was utilized. The suggested design used very little energy and demonstrated significant advantages over the old designs.

Keywords QCA · QCADesigner · T flip-flop

1 Introduction

QCA would be a relatively new class of nanotechnologies which arose in the recent decade. Lent et al. pioneered QCA technique in 1993 [1]. QCA would be a novel approach for computing. QCA would be an attractive alternative to CMOS process for a variety of reasons, including its ultra-high performance, compact size, and power efficiency. It is based on electron configurations rather than input voltages, like CMOS. The primary objective of QCA would be to reduce complexity. Numerous articles have indeed been given on the topic of employing QCA technology to develop significant digital circuitry. The majority of these publications were in quest of the ideal shape. Research has focused on storage cell design as a key circuitry for QCA technologies [2–7], furthermore to a highly efficient flip-flop design [8–11]. The article described an optimal T flip-flop topology with extremely reduced cell count and explained how it may be used to construct extremely efficient *N*-bit counter

S. Gunda (✉) · L. P. Mukkamalla · V. Epuri · P. Pureti · R. S. S. Bathina · S. Bodduboina
Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: sudhakiran.495@gmail.com

circuitry. Numerous counter circuitries were implemented using the provided flip-flop. In QCA technologies, engineers seek to optimize the area, latency, and number of cells needed. As such, optimizing the brick units would be critical. In this article, the chosen domain is represented by a T flip-flop.

The remainder with this work is divided into the following sections: Sect. 2 describes QCA's fundamental concepts, Sect. 3 explains the proposed design, Sect. 4 examines the simulation outcomes, including comparative tables, and Sect. 5 closes this paper.

2 Basics of QCA

2.1 Basic Cell in QCA

QCA circuits have been built from similar components known as QCA cells. As illustrated in Fig. 1, a QCA cell is shaped like a square and has 4 dots at every corner with two oppositely charged electrons that can easily flow through them. Due to the Coulombic force, electrons move toward the dots on diagonally opposed corners. Each cell is somewhat unique due to the fact that it has the right amount of charge. The quantum dot seems to be a nanometer-sized fragment of semiconducting nanostructure with a diameter of between two and ten nanometers. Valence electrons may really migrate to new quantum dots by tunneling via the QCA cell [12, 13].

Coulombic processes may result in the formation of two different polarizations. Each cell keeps two more free electrons in antipodal regions due to Coulombic forces. As a result, the suitable states '0' and '1' have two alternative ground states with linear polarization values of '1' and '− 1.'

The inverter as well as the majority gate constitutes two key components of any QCA circuit. Indeed, QCA's fundamental logic gates contain an inverter and a three-gate configuration. The concept of an inverter is shown in Fig. 2. The inverter has the capability of inverting the number of input signals. In broad strokes, Fig. 3 depicts a three-majority gate. All logic operations in QCA technology are conducted

Fig. 1 Basic QCA cell
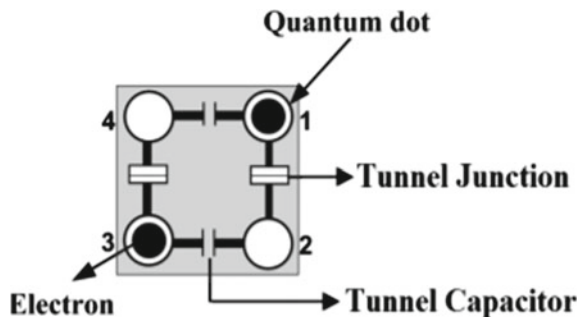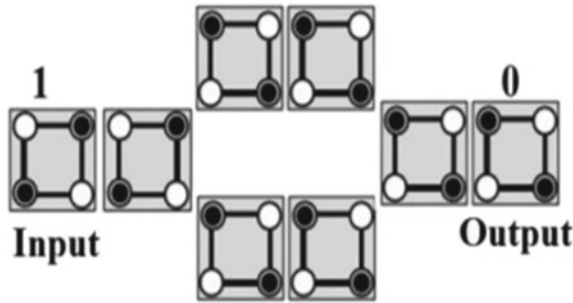
Fig. 2 QCA inverter
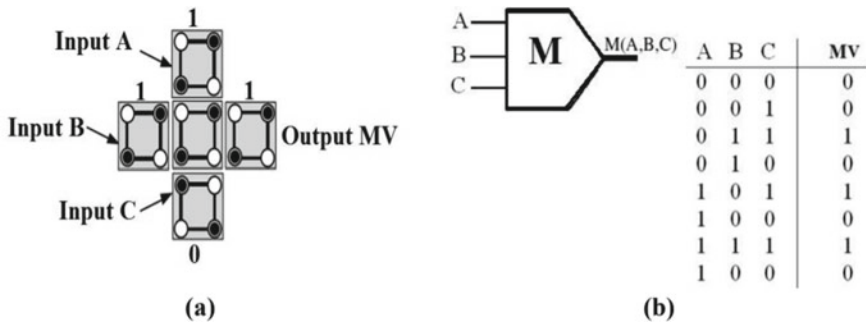
Fig. 3 Majority three-input gate

using an inverter and a three-majority gate. Whether the inverter gate or the three-majority gate was previously used as a critical gate in circuit design, investigations have found that the three-majority gate is underutilized during technology mapping. Certain researchers are still researching novel logic gates in an effort to improve circuit efficiency [14, 15].

The clock is crucial in cellular automata for one variety of reasons: A clock signal synchronizes the circuitry that controls the direction of data flow. Clocking may very well be regarded of as the QCA circuit's principal power source. QCA circuitry achieves intrinsic parallelization by the use of four clocking pulse, each one with four phases, as illustrated in Fig. 4.

3 QCA Implementation of Proposed T Flip-Flop

This section provides unveil a novel TFF structure and afterward demonstrates how to use the relevant converters. The flip-flop shown here would be a level-sensitive model.

Inter-dot Barrier
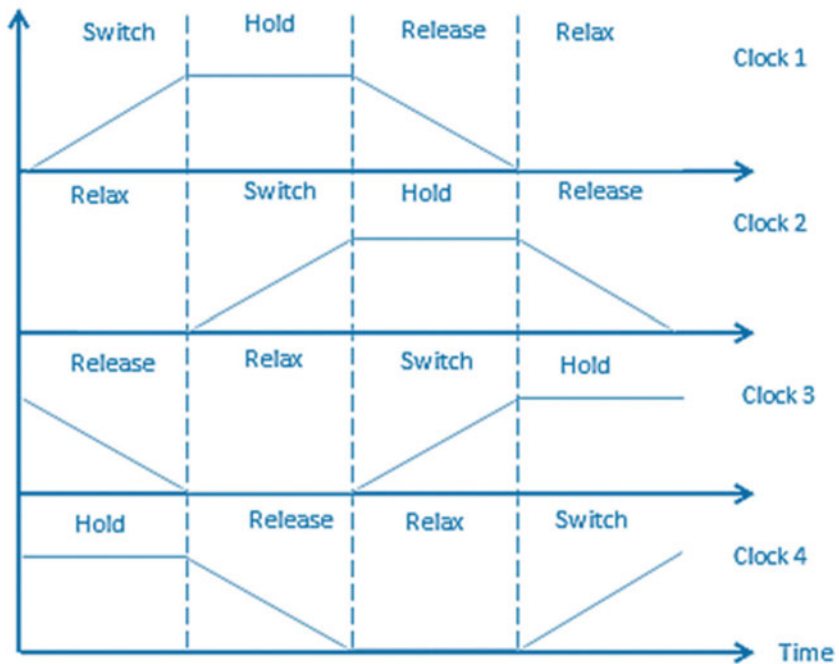
Fig. 4 Clock signals in QCA

The suggested T flip-schematic flop's design is shown in Fig. 5, whereas the QCA arrangement is given in Fig. 6. As seen in Fig. 6, the T flip-flop developed is created by combining the XOR and AND gates. A XOR gate configuration employed in this work is seen in Fig. 7. A circuit approach was used to store the data. When the clock gets available at a high rate as well as the input (T) approaches 1, the output is reversed. As a consequence, no change would have been made to an outcome. The recommended table for the flip-truth flop is shown in Table 1.

Fig. 5 Schematic representation of proposed TFF
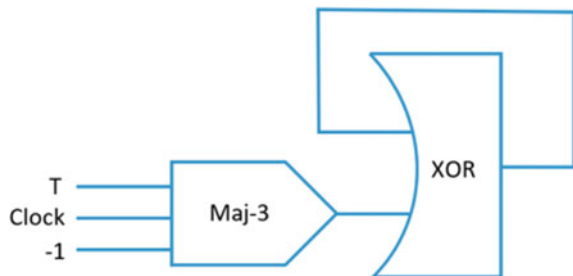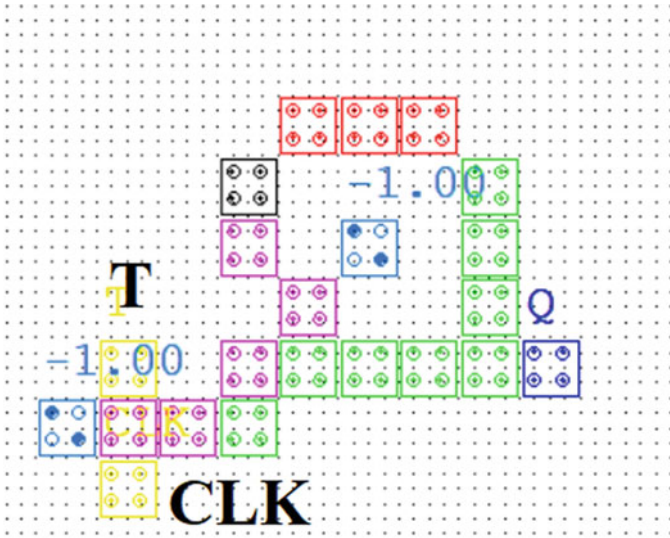
Fig. 6 Proposed TFF
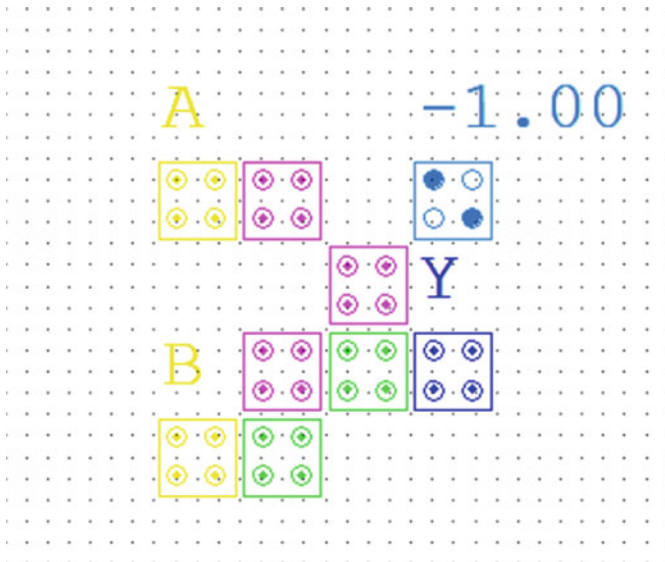
Fig. 7 XOR gate used in TFF

Table 1 Functionality of TFF

T	Clock	Current output
0	0	Q_{t-1}
0	1	Q_{t-1}
1	0	Q_{t-1}
1	1	Q_{t-1}

4 Simulation Results

The suggested structure's effectiveness has been evaluated in this part by contrasting its cell count and area consumption to that of current designs. The result of simulations is performed in QCADesigner version 2.0.3. This work offers a unique T flip-flop architecture that minimizes both cell count and area usage. Figures 8 and 9 illustrate the simulation results for the QCA-based T flip-flop as well as XOR gates, respectively. From Table 2, it is evident that the proposed TFF exhibits better performance.
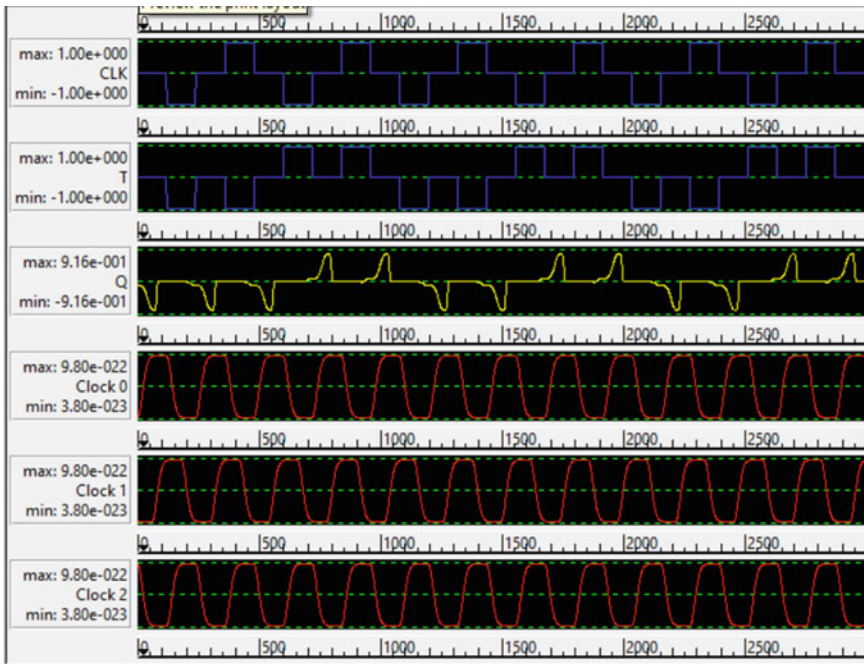
Fig. 8 Simulation result of the QCA-based T flip-flop
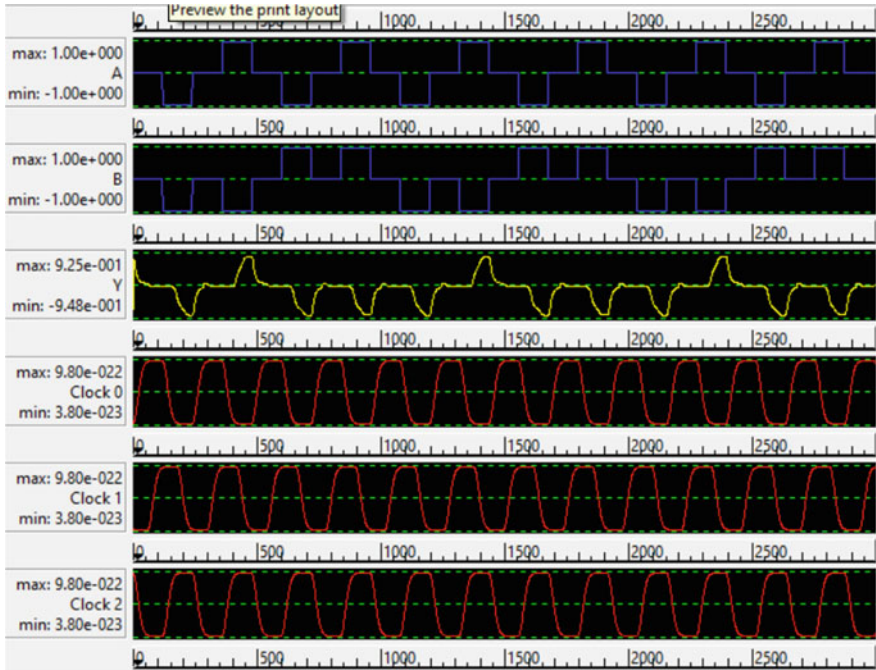
Fig. 9 Simulation result of the QCA-based XOR gate

Table 2 Performance comparison of proposed TFF with others

Design type	Area in (μm^2)	Cell count
[16]	0.06	47
[17]	0.08	67
[18]	0.03	42
Proposed	0.03	22

5 Conclusion

In contrast to CMOS technology, QCA has been nearing nanotechnology having tremendous prospects for providing small circuitry with low energy dissipation. This study describes a small negative-edge triggered TFF. It uses an XOR gate as well as a majority gate in same way as a rising block does, but it requires less space and power. The developed circuitry has been replicated and confirmed using the QCADesigner 2.0.3 tool. This study presented equivalence between many T flip-flops, cell count, and area. The comparison analysis demonstrates that the developed QCA circuit outperforms existing ones.

References

1. Lent CS, Tougaw PD, Porod W, Bernstein GH (1993) Quantum cellular automata. Nanotechnology 4:49–57
2. Dehkordi MA, Shamsabadi AS, Ghahfarokhi BS, Vafaei A (2011) Novel RAM cell designs based on inherent capabilities of quantum-dot cellular automata. Microelectron J 42:701–708
3. Hashemi S, Navi K (2012) New robust QCA D flip flop and memory structures. Microelectron J 43:929–940
4. Angizi S, Sarmadi S, Sayedsalehi S, Navi K (2015) Design and evaluation of new majority gate-based RAM cell in quantum-dot cellular automata. Microelectron J 46:43–51
5. Walus K, Vetteth A, Jullien GA, Dimitrov VS (2002) RAM design using quantum-dot cellular automata. Nano Technol Conf 2:160–163
6. Khosroshahy MB, Moaiyeri MH, Navi K, Bagherzadeh N (2017) An energy and cost efficient majority-based RAM cell in quantum-dot cellular automata. Results Phys 7:3543–3551
7. Asfestani MN, Heikalabad SR (2017) A novel multiplexer-based structure for random access memory cell in quantum-dot cellular automata. Phys B Condens Matter 521:162–167
8. Bhavani KS, Alinvinisha V (2015) Utilization of QCA based T flip flop to design counters. In: Proceedings of the 2015 international conference on innovations in information, embedded and communication systems (ICIIECS), Coimbatore, India, 19–20 March 2015, pp 1–6, 9
9. Yang X, Cai L, Zhao X, Zhang N (2010) Design and simulation of sequential circuits in quantum-dot cellular automata: falling edge-triggered flip-flop and counter study. Microelectron J 41:56–63
10. Angizi S, Moaiyeri MH, Farrokhi S, Navi K, Bagherzadeh N (2015) Designing quantum-dot cellular automata counters with energy consumption analysis. Microprocess Microsyst 39:512–520
11. Sheikhfaal S, Navi K, Angizi S, Navin AH (2015) Designing high speed sequential circuits by quantum-dot cellular automata: memory cell and counter study. Quantum Matter 4:190–197
12. Roshany HR, Rezai A (2019) Novel efficient circuit design for multilayer QCA RCA. Int J Theor Phys 58(6):1745–1757
13. Balali M, Rezai A, Balali H, Rabiei F, Emadi S (2017) Towards coplanar quantum-dot cellular automata adders based on efficient three-input XOR gate. Results Phys 7:1389–1395
14. Das JC, De D, Mondal SP, Ahmadian A, Ghaemi F, Senu N (2019) QCA based error detection circuit for nano communication network. IEEE Access 7:67355–67366
15. Chabi AM, Roohi A, DeMara RF, Angizi S, Navi K, Khademolhosseini H (2015) Cost-efficient QCA reversible combinational circuits based on a new reversible gate. In: 18th CSI international symposium on computer architecture and digital systems (CADS), pp 1–6
16. Angizi S, Moaiyeri MH, Farrokhi S, Navi K, Bagherzadeh N (2015) Designing quantum-dot cellular automata counters with energy consumption analysis. Microprocess Microsyst 39(7):512–520
17. Surya Bhavani K, Alinvinisha V (2015) Utilization of qca based t flip flop to design counters. In: 2015 international conference on innovations in information, embedded and communication systems (ICIIECS). IEEE, pp 1–6
18. Karthik R, Kassa SR (2018) Implementation of flip flops using qca tool. J Fundam Appl Sci 10(6S):2332–2341

Efficient Image Watermarking Using Particle Swarm Optimization and Convolutional Neural Network

Manish Rai, Sachin Goyal, and Mahesh Pawar

Abstract In today's Internet environment, copyright protection of digital multimedia material is a major concern. Digital image watermarking provides the concept of secured copyright protection—the strength and quality of the watermark image loss due to geometrical attacks. However, minimization of the impact of geometrical attacks is a major challenge. This paper proposed a coefficient optimization-based watermarking algorithm. The particle swarm optimization applies to feature optimization of the source and symbol images. The processing of feature extraction of the source image and symbol image uses the wavelet transform function. The CNN algorithm follows the process of embedding with an optimal coefficient of the PSO algorithm. The proposed algorithm has been tested on MATLAB environments with reputed image datasets. The performance of the proposed algorithm is estimated as correlation coefficient and PSNR. The estimated results compare CNN algorithm and WCFOA. The improved outcome of the proposed algorithm against the existing algorithm was approx. 2%.

Keywords Image watermarking · DWT · PSO · CNN · Geometrical attacks · Optimization

1 Introduction

The copyright protection of digital multimedia is a significant concern in the current scenario of Internet technology. The security of intellectual property rights is provided by digital watermarking. Digital watermarking applies to all types of multimedia data, including images, videos, audio, and text. Images make up the majority

M. Rai (✉) · S. Goyal · M. Pawar
RGPV University, Bhopal, India
e-mail: manishrai2587@gmail.com

S. Goyal
e-mail: sachingoyal@rgtu.net

M. Pawar
e-mail: mkpawar24@gmail.com

of digital multimedia data. 70% of the data generated and sent over the Internet is covered [1, 2]. The utility of digital image watermarking has increased in recent years by printing currency and legal documents. However, despite several watermarking algorithms, the quality and robustness of watermarking are still a big challenge. The research of digital image watermarking is divided into two domains: spatial and frequency. The security potential of watermarking methods was revealed in the spatial domain due to algorithm limitations. On the other hand, the frequency domain of the watermarking algorithm provides much security strength of watermarking [3–6].

Most watermarking methods are based on transform functions such as discrete wavelet transform, DCT, WPT, and many derivatives of wavelet transform function. Still, the strength of the wavelet transforms-based watermarking algorithm is compromised with geometrical attacks such as rotation, scaling, and translation. The feature optimization-based watermarking is the new direction of robust watermarking against geometrical attacks. There are several criteria in image watermarking, including the host image not being corrupted after embedding the image, the watermark being imperceptible from the watermarked image, the watermark being difficult to damage or remove, and the watermark being highly robust from the host image. Watermarks, on the other hand, are resistant to a variety of attacks, including scaling noise, blurring, filtering, rotation, JPEG compression, and cropping. Watermarks are highly secure and difficult for unauthorised users to access [7]. The feature optimization-based watermarking methods reduce the gaps of correlation coefficient and increase the strength of the watermark image.

The proposed algorithm integrates particle swarm optimization and the CNN model for embedding the watermark [8–10]. The layering of CNN moves on the complex structure of the watermark image and resists geometrical attacks. The processing of the proposed algorithm encapsulates several phases, feature extraction, feature transformation, feature optimization, coefficient selection, embedding and applying attacks to measure correlation coefficient, and PSNR parameters [11]. The significant contribution of this paper is the proposed efficient watermark method for multimedia data. The next goal of this paper is to test the suggested strategy with geometrical attacks and determine the correlation coefficient (CC). The main objective of this work is [11] to propose efficient image watermarking methods using feature optimization and the CNN algorithm. Furthermore, [12] enhances the visual perception of watermark images in terms of PSNR. Finally, tested the robustness factors on different types of attacks. The remainder of the work is laid out as follows. The second half of the study covers well-known image watermarking techniques. The proposed approach is explained in Sect. 3, followed by an experiment evaluation in Sect. 4, and finally a judgement in Sect. 5.

2 Related Work

The incremental approach of watermarking algorithms resolves many copyrights protection and ownership issues. The feature-based watermarking methods provide

good results and security with specific limitations. Some contributions of authors describe here.

In [13], authors used an ensemble-based algorithm and their proposed algorithm produce a better watermark image against different types of attack.

In [14], authors give a brief survey about different types of watermarking methods against different types of geometric attack.

In [1], authors concluded that instead of establishing a neural network to protect a single image, their form of watermarking uses DNNs. They also went through some of the trained model's characteristics that make it vulnerable to JPEG compression attacks.

In [2], authors created a new way to copyright colour photographs without needing to search at them the proposed approach, which is based on crossing inter-block quaternion discrete wavelet transform, incorporates extreme pixel rectification, multi-bit partially key indicator means modulation, mixed measurement, and particle swarm optimization.

In [3], authors proposed a safe watermarking approach that improves invisibility and robustness at the same time. Particle swarm optimization and the firefly algorithm are used. The PSNR, NC, NPCR, and UACI values are 55.1471 dB, 0.9956, 0.9801, and 0.3471, respectively.

In [5], authors ensure tamper detection and authenticity, and the twin water-marks were buried in the area of non-interest block of the MRI pictures. Simulation studies on several types of medical images reveal that the suggested methodology has improved transparency and robustness against signal and compression attacks when compared to earlier hybrid optimised methodologies.

In [6], authors propose a perceptron for blind steganographic digital images that is both robust and clear. It is a CNN-based system with pre-processing nets for both the host picture and the watermark, as well as a watermark embedding network, a training simulation, and a background subtraction network to extract the watermark as needed.

In [7], authors propose a combination of LWT and SVD, as well as the TMR methodology, and are offered as a durable and reliable digital picture watermarking method. According to the assessments, the presented method is sturdy and has a high imperceptibility when compared to existing works.

In [15], authors propose that the RCNN is a mapping-based residual convolution neural network. It is a blind photo watermarking solution that does not require any embedding. According to the findings of the testing, the suggested system can successfully extract the watermark photographs under various attacks and is even more robust than existing content filtering methods.

In [16], authors propose that WCFOA is an image watermarking system based on deep CNN that is developed with the goal of inserting a secret message in the cover image. According to the findings, utilising measurements like correlation and PSNR and the WCFOA-based deep CNN, improved its performance, with values of 1 and 45.22 for that one without noise scenario and 0.98 and 45.08 for the various noise situation.

In [17], authors create an ideal clustering approach to create an efficient digital watermarking system phases such as these are included in the proposed model. The experimental findings will lead to a higher demand for copyrighting in term of resilience and responsiveness.

In [18], authors collect the plant–village dataset, which can be downloaded for free. A hybrid basis of ratio assessment optimization technique is used to extract useful features from the dataset.

In [8], authors devised a method that starts by using a SMFMF to extract the noise-affected pixels from a clinical image, then replacing the noisy pixel with the median pixel's value. The PSNR and normalised cross correlation values are used to assess the clinical image's multi-target wellness forecasts. The results suggest that the proposed approach, which uses watermarking validation, improves clinical strength and protection.

In [9], authors presented the overall idea of a watermarking framework is examined using a neural network-based approach to deal with colour images under a range of attacks. Using the innovative integration of wavelet transform in combination with the multilayer perceptron to extract the logo information, a variety of techniques is examined to deal with the above-mentioned watermarking framework.

In [10], authors proposed the IoT perceiving layer, and the DSRW method is utilised, which is based on a hybrid image security optimization algorithm in the DWT domain.

In [19], authors proposed a learning and neural network-based supply chain risk management model. Furthermore, based on the current state of supply chain management, this paper examines the risk indicator system.

In [20], authors introduce a unique zero-bit watermarking approach. This method allows the watermark to be extracted in the future with only a few queries while preserving the protected model's performance.

In [21], authors use a blind image steganographic strategic plan based on PCA in the redundant discrete wavelet domain that is also streamlined throughout terms of quality measures like the normalised correlation coefficient, peak signal-to-noise ratio, mean square error, and structural similarity index; the proposed framework is resilient.

In [22], authors proposed the LWT and DNN, a robust image watermarking method. The outcomes reveal that the suggested method outperforms state-of-the-art techniques in practically all parameters for the vast majority of situations.

In [23], authors describe quantum neural networks' machine learning dynamics which are studied, as well as quantum chaos and complexity. The outcomes show that the generalization ability of QNN is maximised whenever the system is operating system or oscillation in phase space.

In [24], authors proposed a blind and robust method using the YCBCR colour space, IWT, and DCT was created for colour image watermarking. In terms of imperceptibility and resilience, the results demonstrate that they are superior.

In [11], authors proposed DNN model for digital watermarking, intellectual property, embedding watermarks, and owner verification are all investigated. The result

establishes property of all remotely expanded deep learning models in a reliable and timely way, while also ensuring model accuracy for standard data.

In [12], author describe a new digital watermarking technique for deep learning models that produce images, in which any image produced by a watermark information neural net must include a specific signature.

3 Proposed Methodology

The proposed algorithm is executed in three phases; in the first phase process, the feature extraction using discrete wavelet transform; in the second phase, estimate optimal position coefficient for the embedding and finally describe the embedding approach of the watermark based on CNN network.

3.1 Feature Extraction

Feature extraction is the primary phase of image watermarking. The feature extraction process applies discrete wavelet transform (DWT); the applied transform decomposes the raw images in multiple forms of LF HF horizontal, vertical, and diagonal. The decompose layers estimate the value of features in terms of energy entropy [8].

Let $L(x) \in L^2(I)$ is correlated function (x) and scaling function $\phi(x)$, the formulation of transform as

$$W_\phi(j0, k) = \frac{1}{\sqrt{M}} \sum_x L(x)\phi_{j0,k}(x) \tag{1}$$

$$W_\psi(J, K) = \frac{2}{\sqrt{M}} \sum_x L(x)\psi_{j,k}(x) \tag{2}$$

The coefficient of feature of images as

$$F(L) = \frac{1}{M} \sum_K W_\phi(j0, k)\phi_{j0,k}(x) + \frac{1}{\sqrt{M}} \sum_{J=J0}^{J-1} \sum_K W_{\psi(J,K)}\psi_{J,K}(X) \tag{3}$$

The above derivatives of function estimate the feature of raw image in terms of source image and symbol image. The estimated features

$$F(L) = \{f1, f2, f3, \ldots, fn\} \tag{4}$$

3.2 Optimization Coefficient (K)

The optimization coefficient is used to choose the best placements for symbol pictures to be embedded in the source image. With particle swarm optimization, the procedure of coefficient optimal control estimates. A well-known swarm intelligence algorithm is particle swarm optimization. The acceleration function, constant factor, initial velocity, ultimate velocity, and particle location are used to process particle swarm optimization. The algorithm's processing is described here [20].

The description of optimization of coefficient describes as S-feature space vector of particle. Now, the space of particle $xi = (xi1, xi2, xi3, \ldots, XiS)$. This represents of vector of velocity space as $Vi = (Vi1, Vi2, \ldots, Vis)$. The processing of features changes the position of particle vector $Pi = (pi1, pi2, \ldots, pis)$. The estimation of updated position as

$$vis = Vis + C1a1(Pis - xis) + c2a2(Pgs - xis) \tag{5}$$

$$xis = xis + Vis \quad s = 1, \ldots s, \ldots S \tag{6}$$

the range of factor is [0,1]

the selection coefficient as $K = C1 + C2$

$$k = \begin{cases} \frac{2x}{L(p)-2+\sqrt{x^2-4x}} & \text{for } X > 4 \\ x, & \text{otherwsie} \end{cases} \tag{7}$$

The value of K updated the region of feature coefficient

$$Vis = K(vis + c1a1(Pis - xis) + c2a2(Pgs - xis)) \tag{8}$$

3.3 Embedding of Watermark

After the feature extraction, the grid matrix of features of the source image and symbol image applies the CNN model for embedding. The CNN model work along with the optimization coefficient K for the optimal embedding position of the watermark image. The formation of the algorithm describes here.

The processing of CNN features procedures and finds the optimal position for more effectively watermark embedding. The process of CNN proceeds as convolution layer, pooling layer, and fully connected layer. Each layer of the CNN network performs a distinctive function for constructing features maps. The pooling layers deal with the sampling of features, and the fully connected layer estimates the embedding position.

Convolution layer—the potential of convolution to derives the patterns of features using the filter.

The output of filter estimated as fy.

$$fy = \left(S_q^P\right) + \sum_{g=1}^{K} \sum_{\delta=k}^{k1} \sum_{h=k1}^{k2} \left(F_q^p\right) * \left(f_g^{q-1}\right)g + \delta + k \tag{9}$$

Here, $*$ is convolution and k is features of filters and variable g, K is range of maps of features. Pool layer: the processing of pooling layer deals with fixed operator operation. Fully connected layers: the output of pooling layers' process as input of FC layer. After the processing of layers, estimated the final region for the embedding. Terms define for the process of embedding of watermarking.

SI = source image, Symbol image = Sm, Reu = learning factor of CNN, FF = fully connected, ROI = region of interest, K = coefficient of optimization. Tr = training rate of model Wm = watermark image (Fig. 1).

Algorithm 1. Embedding process

Input: set parameters (SI, Sm, K)

Output: $O - $ Wm (watermark image)

Begin

Step 1. Mapped all extracted features in CNN
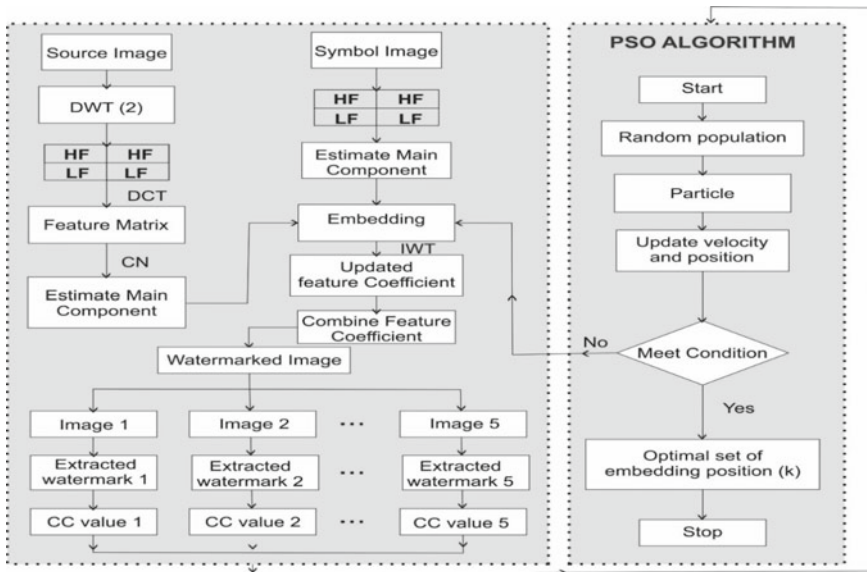
Fig. 1 Proposed model of watermarking approach using PSO and CNN

Step 2. Feature matrix transform into polling factors

Step 3. Training $Tr(Rlu; Xi)$ for each layer of features, $FFi = 1, \ldots, n$

Step 3. $k = FFi$; set of connect feature of ROI

$$Wm \leftarrow FF\{Tr\}; \; SI \leftarrow K\{xi\}; \; MFFi = FFi - 1$$

Step 4. while $Wm \neq \emptyset$ **do**

    Change the position of symbol $Smi \in \{1, 2, \ldots, X\}$;

$$Sm = sm - 1;$$

    if $(whight > 0)$ **then**

$Wm \leftarrow SIf \cup \{W_i\}$

    end

end

Step 5 embedding is done

Return Wm

4 Experimental Analysis

The validation of proposed watermarking algorithm simulates in MATLAB tools using five source image and five symbol images. All 10 images resolution size is 512×512 and 256×256 with dip 32. The processing of training and coefficient optimization is written in MATLAB function and directly applied on watermark process. The experimentation configuration of system is I7 processors, 16 GB ram, and windows operating system. The analysis of results parameters estimated as correlation coefficient and peak signal noise ratio (PSNR) [16].

$$CC = \frac{WM}{SISm} \tag{10}$$

$$PSNR = 20 \log 10 \frac{W \max X \times XY}{\sum \sum (Mw(x, t) - Si(x, t))} \tag{11}$$

Performance of watermark reliability for embedding colour watermarks using CC and PSNR dataset is determined between the images generated by WCFOA, CNN, and proposed method, and results are shown in Table 1.

Table 1 Performance assessment of watermark reliability by WCFOA, CNN, and proposed method

S. No.	Method	Host	Symbol	Attack	CC	PSNR
1	WCFOA [17]	Boat	VO	Row-column blanking	0.937	43.578
2	CNN [16]				0.951	43.781
3	Proposed				0.882	44.359
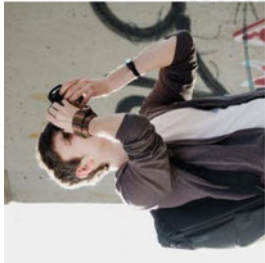
HOST IMAGE   SYBOL IMAGE   WATERMARKED IMAGE
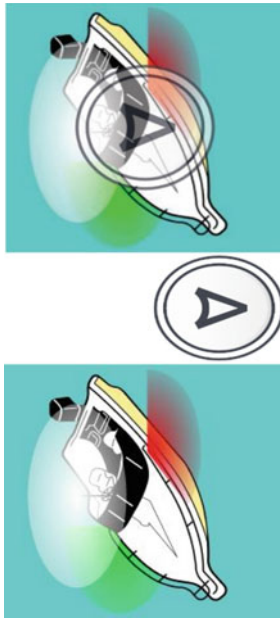
4	WCFOA [17]	Cameraman	Nike	Scaling	0.967	44.197
5	CNN [16]				0.953	44.217
6	Proposed				0.921	45.625

(continued)

Table 1 (continued)

S. No.	Method	Host	Symbol	Attack	CC	PSNR
		HOST IMAGE	SYBOL IMAGE	WATERMARKED IMAGE		
		Crane	Crown	Translation		
7	WCFOA [17]				0.784	46.283
8	CNN [16]				0.779	46.952
9	Proposed				0.771	47.154

(continued)

Table 1 (continued)

S. No.	Method	Host	Symbol	Attack	CC	PSNR
		HOST IMAGE	*SYBOL IMAGE*	*WATERMARKED IMAGE*		
10	WCFOA [17]	Tiger	Skull	Rotation	0.847	43.743
11	CNN [16]				0.814	43.965
12	Proposed				0.807	44.796

(continued)

Table 1 (continued)

HOST IMAGE   SYBOL IMAGE   WATERMARKED IMAGE

S. No.	Method	Host	Symbol	Attack	CC	PSNR
13	WCFOA [17]	Home	Square	Cropping	0.916	47.351
14	CNN [16]				0.904	47.893
15	Proposed				0.895	48.654

(continued)

Table 1 (continued)

S. No.	Method	Host	Symbol	Attack	CC	PSNR

HOST IMAGE

SYBOL IMAGE

WATERMARKED IMAGE

Fig. 2 CC analysis for different host images
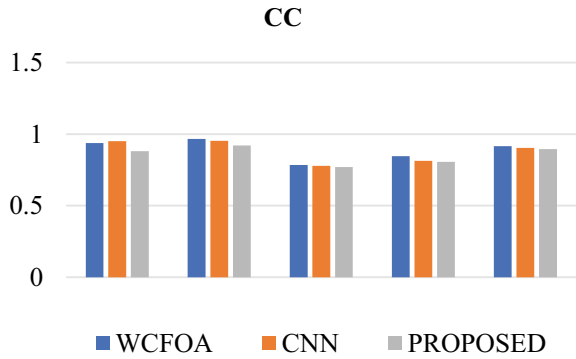
Fig. 3 PSNR analysis for different host images
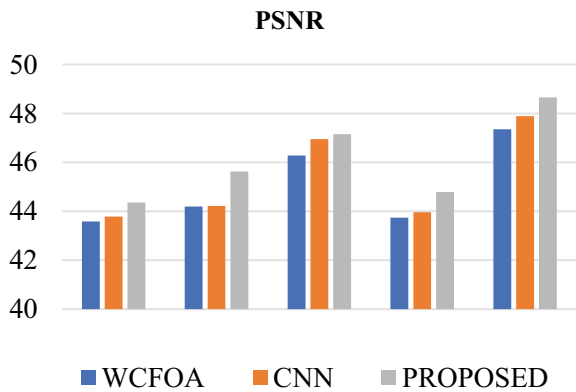
Figures 2 and 3 show analysis of watermark, CC, and PSNR values with different images using WCFOA, CNN, and proposed method of different host, boat, cameraman, crane, tiger, and home dataset.

5   Conclusion and Future Work

This paper proposed an efficient digital image watermarking method using coefficient optimization and CNN algorithms. The proposed algorithm estimates the optimal position of embedding of watermark coefficient K. The iteration of particle swarm optimization finds the best coefficient value K for embedding. The applied CNN model compacts the embedding process using Rlu learning factors. The security strength of the proposed algorithm was estimated by applying different geometrical attacks such as rotation, scaling, and others. It calculated the PSNR and CC values. The low value of CC demonstrates the watermark algorithm's potential power. The PSNR value is used to determine the visual quality of the watermark image.

PSNR's assessed findings indicate that the watermark image's quality is unaffected. The results in Table 1 show the different attacks and consequences of paraments. The proposed algorithm compares with CNN and WCFOA algorithms. The evaluation of the results shows the strength of the watermark algorithm. The statistical improvements in CC value are approx. 1–2% with existing algorithms.

References

1. Ding W, Ming Y, Cao Z, Lin C-T (2021) A generalized deep neural network approach for digital watermarking analysis. IEEE Trans Emerg Topics Computat Intell
2. Hsu L-Y, Hwai-Tsu H (2020) Blind watermarking for color images using EMMQ based on QDFT. Expert Syst Appl 149:113225
3. Mohan A, Anand A, Singh AK, Dwivedi R, Kumar B (2021) Selective encryption and optimization-based watermarking for robust transmission of landslide images. Comput Electr Eng 95:107385
4. Singh R, Ashok A (2021) An optimized robust watermarking technique using CKGSA in frequency domain. J Inf Secur Appl 58:102734
5. Swaraja K, Meenakshi K, Kora P (2020) An optimized blind dual medical image watermarking framework for tamper localization and content authentication in secured telemedicine. Biomed Signal Process Control 55:101665
6. Lee J-E, Seo Y-H, Kim D-W (2020) Convolutional neural network-based digital image watermarking adaptive to the resolution of image and watermark. Appl Sci 10(19):6854
7. Salehnia T, Fathi A (2021) Fault tolerance in LWT-SVD based image watermarking systems using three module redundancy technique. Expert Syst Appl 179:115058
8. Balasamy K, Shamia D (2021) Feature extraction-based medical image watermarking using fuzzy-based median filter. IETE J Res, 1–9
9. Kazemi MF, Pourmina MA, Mazinan AH (2020) Analysis of watermarking framework for color image through a neural network-based approach. Complex Intell Syst 6(1):213–220
10. Alam S, Ahmad T, Doja MN, Pal O (2021) Dual secure robust watermarking scheme based on hybrid optimization algorithm for image security. Pers Ubiquitous Comput: 1–13
11. Deeba F, She K, Dharejo F, Memon H (2020) Digital watermarking using deep neural network. Int J Mach Learn Comput 10(2):277–282
12. Wu H, Liu G, Yao Y, Zhang X (2020) Watermarking neural networks with watermarked images. IEEE Trans Circuits Syst Video Technol
13. Rai M, Goyal S, Pawar M (2021) Feature optimization of digital image watermarking using machine learning algorithms. In: Bajpai MK, Kumar Singh K, Giakos G (eds) Machine vision and augmented intelligence—theory and applications. Lecture notes in electrical engineering, vol 796. Springer, Singapore. https://doi.org/10.1007/978-981-16-5078-9_39
14. Rai M, Goyal S, Pawar M (2019) Mitigation of geometrical attack in digital image watermarking using different transform based functions. Int J Innov Technol Exploring Eng (IJITEE) 8(9). ISSN: 2278-3075
15. Wang X, Ma D, Kun H, Jianping H, Ling D (2021) Mapping based residual convolution neural network for non-embedding and blind image watermarking. J Inf Secur Appl 59:102820
16. Ingaleshwar S, Dharwadkar NV, Jayadevappa D (2021) Water chaotic fruit fly optimization-based deep convolutional neural network for image watermarking using wavelet transform. Multimedia Tools Appl: 1–25
17. Soppari K, Subhash Chandra N (2020) Development of improved whale optimization-based FCM clustering for image watermarking. Comput Sci Rev 37:100287
18. Gadekallu TR, Rajput DS, Praveen Kumar Reddy M, Lakshmanna K, Bhattacharya S, Singh S, Jolfaei A, Alazab M (2020) A novel PCA–whale optimization-based deep neural network model for classification of tomato plant diseases using GPU. J Real-Time Image Process 1–14

19. Han C, Zhang Q (2021) Optimization of supply chain efficiency management based on machine learning and neural network. Neural Comput Appl 33:1419–1433
20. Le Merrer E, Perez P, Trédan G (2020) Adversarial frontier stitching for remote neural network watermarking. Neural Comput Appl 32(13):9233–9244
21. Rajani D, Rajesh Kumar P (2020) An optimized blind watermarking scheme based on principal component analysis in redundant discrete wavelet domain. Signal Process 172:107556
22. Mellimi S, Rajput V, Ansari IA, Ahn CW (2021) A fast and efficient image watermarking scheme based on deep neural network. Pattern Recognit Lett 151:222–228
23. Choudhury S, Dutta A, Ray D (2020) Chaos and complexity from quantum neural network: a study with diffusion metric in machine learning. arXiv preprint arXiv:2011.07145
24. Sinhal R, Jain DK, Ansari IA (2021) Machine learning based blind color image watermarking scheme for copyright protection. Pattern Recognit Lett 145:171–177

A Review on Various Cloud-Based Electronic Health Record Maintenance System for COVID-19 Patients

D. Palanivel Rajan, C. N. Ravi, Desa Uma Vishweshwar, and Edem Sureshbabu

Abstract Cloud computing is an evolving technology to maintain the database of any system. Data collected from any part of the system will be transferred to the cloud, and it will be retrieved at any point in time. It plays a vital role in biomedical applications, where a huge number of patient records are needed to be maintained. In recent years, we faced an unexpected pandemic condition due to COVID-19 diseases. Routine human life has turned upside down due to it. This disease affects various age groups of people, and the number of patients affected is also growing exponentially, day after day, across the globe. The treatment for this critical illness is not the same for patients of different age levels. Aged people may be already affected by various diseases, whereas middle-aged and children may not be. COVID-19 is getting more vulnerable, and the death rate is increasing. Diagnosing this disease is a tedious task for doctors. Symptoms collected from patients of various ages and the treatment methods offered to them should be appropriately maintained. This may ease out ways to cure the upcoming affected patients. In this paper, we present an overall review of various cloud-based electronic health care recording methods that are currently available. Personal health records (PHRs) are stored on a remote cloud server using these approaches. The selective information will be shared when needed by a physician for diagnosis and treatment. Many new cloud-based systems are developed, which have more secure and safe data transfer compared to the conventional client–server model. Security is the most concerned parameter for the emerging cloud technologies as PHRs are to be maintained confidentially. The various existing cloud-based models are reviewed in the aforementioned aspect.

Keywords COVID-19 · Electronic health record · Cloud · Patient health record · Security

D. Palanivel Rajan (✉) · C. N. Ravi · D. U. Vishweshwar · E. Sureshbabu
Department of CSE, CMR Engineering College, Hyderabad, Telangana, India
e-mail: palanivelrajan.d@gmail.com

C. N. Ravi
e-mail: cnravi@cmrec.ac.in

E. Sureshbabu
e-mail: sureshbabu.edem@cmrec.ac.in

1 Introduction

The world is in the latest situation with the outbreak of coronavirus disease 2019. Since, December 2019, this disease is causing serious impairment and threats in more than 200 countries worldwide. This condition also reveals that many patients are asymptomatic, or only have mild symptoms [1]. Patients with asymptomatic symptoms caused by COVID-19 have recovered by maintaining quarantined and taking medicines in some cases, however, the incidence of mortality has increased due to patients suffering from serious coronavirus disease with coagulopathy [2], and often COVID-19 also becomes problematic for patients suffering from acute pulmonary embolism and right-sided heart failure [3].

Doctors in different parts of the world need to monitor the people affected with coronavirus and also need to provide care to them in different countries. For this, the physician needs to maintain the health record of a patient with a history of whether he/she has already been affected by any other diseases or only affected by this disease. To adopt this method, physician needs to create an electronic health record (EHR) of the patient with incorporate cloud computing technology for easy access of a record from anywhere in the world by the physician, EHR system helps to the physicians to monitor the flow of patient files inside a hospital and also retrieve the information about the diseases around the world and provide specific care to the patient's for getting better lifecycle quality to maximize data, and recovery processes [4].

At the same time, records that are stored in cloud-based technology need to be maintained in a secure way, and also authenticated persons only access that data whenever they need it. In order to efficiently store and exchange PHRs and eradicate patients' concerns about PHR privacy, the secure cloud-based EHR approach, which ensures the protection and security of data that is stored in the cloud, is discussed using various techniques to rely on cryptographic primitives.

In Sect. 2 reviews corona diseases, different types of symptoms, how they spread to each other, a comparison of diseases in various countries like developing and developed countries, and treatment is given for a disease. In Sect. 3 discusses the importance of electronic health record, how it was designed and implemented in various countries with help of cloud computing technology, and also the various security concern mechanism is used to maintain the data in cloud computing are discussed.

2 Corona Virus

Coronavirus 2019 is a transferrable disease instigated by extreme ARS (SARS- CoV-2). This disease is identified in Wuhan, China in December 2019 and now it has resulted in a current epidemic. As of now, over 26.6 million in 188 countries and territories have been registered with over 874,000 deaths, with over 17.7 million recovered. As in Fig. 1, various possibilities like fever, cough, fatigue, shortness of
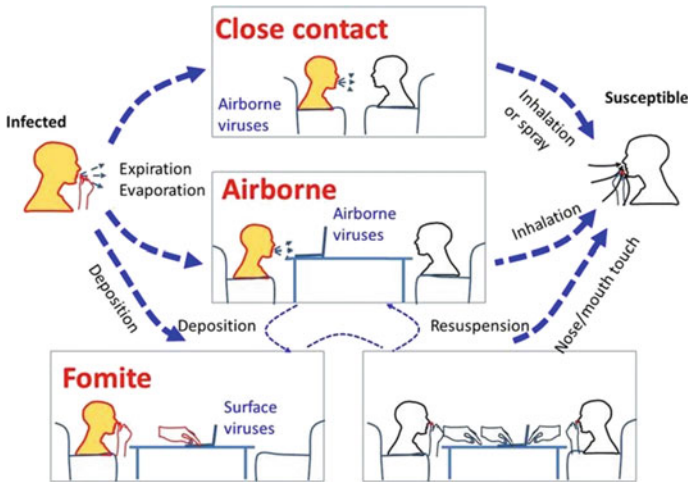
Fig. 1 Possibilities of a cause of infection

breath, and smell and taste issues are among the most common symptoms. While most people have mild symptoms, others have ARDS syndrome, which is precipitated by a cytokine storm.

2.1 Possibilities of Transmission

This type of disease that spread is heavily influenced by age as four ranges have been reported: 0 to 14, 15 to 44, 45 to 64, and $\implies$ 65 years [5]. Figure 2 shows the age-wise distribution of cases affected in Mumbai. Mainly by small coughing, sneezing, and talking droplets the virus is transmitted. Usually, the droplets are not airborne. However, those in the immediate vicinity can inhale and get contagious. People may get contaminated by rubbing and then pressing on their mouths. The transfer can also be carried out by tiny droplets, which can be suspended for longer periods in the air in closed spaces. There are three possibilities of transmission as symptomatic, pre-symptomatic, and asymptomatic where people can infect.

Symptomatic Transmission. The term "symptomatic transmission" refers to transmission from a person who has symptoms. This type of virus is transmitted from patients who are infected already to people in near distance by respiratory droplets, contact with polluted objects and surfaces, and straight contact with infected persons.

Pre-symptomatic Transmission. The interval between being exposed to this type of virus (getting infected) is on average 5–6 days but can be up to 14 days. Some infected people can be contagious during this phase, which is also known as the

Fig. 2 Age-wise distribution

"PR symptomatic" period. As a result, transmission from a pre-symptomatic illness might occur prior to the emergence of symptoms.

Asymptomatic transmission. In this, a virus is transmitted from one person to another, who does not have any symptoms. This disease is the most contagious during the first three days after the beginning stage of symptoms, but it can spread before and in people who are asymptomatic.

2.2 Diagnosis Methods

A nasopharyngeal swab is the basic diagnostic tool by reverse transcription-polymerase (RT-PCR). Chest CT imagery will benefit people with elevated fears of infection based on symptoms and risk factors to be diagnosed.

2.3 Prevention Methods

Preventing infection risk minimizations includes staying in a house, wearing a mask in public places, maintaining a distance from others, avoiding crowded places, regularly washing the hands for at least 20 s with soap and water, proper respiratory care, and avoid of touching eyes, nose, or mouth without washing your hands. Let us see some reviews on this type of covid19 attack.

2.4 Severe Coronavirus Disease Patients with Coagulopathy

Coagulopathy is a common complication of severe coronavirus disease where disseminated intravascular coagulation (DIC) may cause more number of deaths. Many patients with severe diseases may be affected due to Third International Consensus Definitions for Sepsis for the evidence of virus infection and respiratory dysfunction (Sepsis-3). In addition to these severe cases long-term bed rest and hormone treatment raise the risk of venous thromboembolism (VTE).

Due to these reasons, it is recommended that individuals with severe can take anticoagulants (such as heparin) on a regular basis. The International Society of Thrombosis and Hemostasis proposed a new approach known as "sepsis-induced coagulopathy" (SIC) to analyze an earlier stage of sepsis-associated DIC, and it has been confirmed that patients who meet the diagnostic criteria for SIC benefit from anticoagulant therapy. They used retrospective analysis for validating the effectiveness of the SIC score and other coagulation indicators in the screening of outpatients who could get benefit from anticoagulants.

3 Implementation of Electronic Health Record

The implementation of electronic health records with cloud computing technology is simple and easy. It is a wide technology that provides less cost for maintaining records, security, privacy, scalability, and implementation. As listed in Table 1, it is platform-independent with reduced errors, improved quality, flexibility, and exchange aid sharing ability. Owing to these features, cloud computing technology can effectively provide the implementation support for the electronic health record as shown in Fig. 3 [6].

Moreover, electronic health record is the digital version of patient medical history which overcomes the record maintaining in paper. The main objective of creating and maintaining EHR is it helps physicians to make a decision and disease diagnoses for the patients [7]. When he arrives with any disease, the physician needs to check the previous records.

3.1 Attribute-Based Encryption

It is otherwise known as fuzzy identity-based encryption. Key-policy attribute-based encryption (KP-ABE) and ciphertext-policy attribute-based encryption (CP-ABE) were defined as two different and complementing conceptions of ABE in this method (CP-ABE). A construction of KP-ABE was provided in the same paper [8, 9], while the first CP-APE construction supported for tree-based access method in a generic group model. Subsequently, a number of alternatives of ABE schemes have been

Table 1 Advantages of EHR in the cloud

S. No.	Domain	Uses
1	Cost	Reducing software, hardware, installation, and maintenance costs
2	Security and Privacy	Encryption and decryption, Authentication, Two- factor authentication, Confidentiality, Audit
3	Scalability	Provision of resources at the time of demand, the Possibility of development
4	Implementation on	PC, Laptop, Mobile, Windows, SQL, ORACLE, open EHR
5	Flexibility	Use of multiple and separate sites, use for all shareholders
6	Exchange and sharing ability	Fast transfer, time-dependent, the permission of file owners
7	Reducing errors and improving the quality	Reducing takes in exchanges, increasing access speed, integrating the multiple types of electronic health records
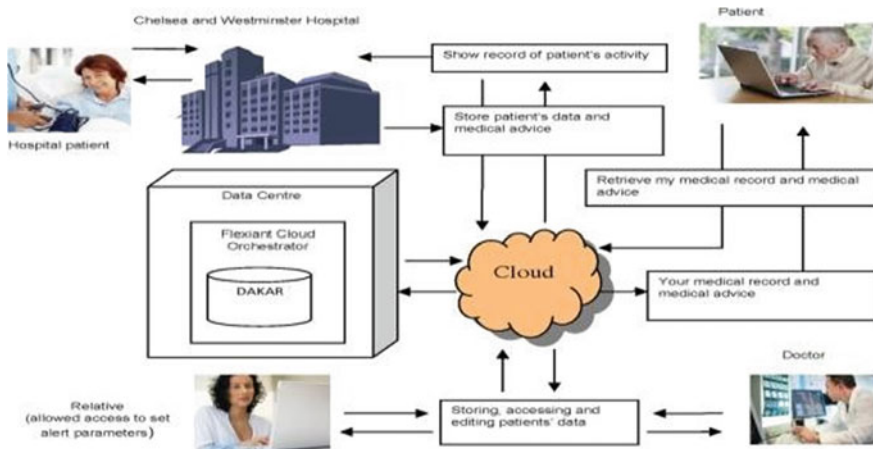
Fig. 3 Accessing a record in the cloud

proposed since its introduction. Its features are being expanded, and schemes with stronger security are being proposed. This scheme is an example, of accepting any type of access structure. The common thing in this approach is to use a technique to delegate overhead computing with secure outsourcing for encryption/decryption/key-issuing to the third parties, which reduces the amount of the calculation done locally. This strategy aids us to develop an outsourcing paradigm for achieving efficient local decryption [10].
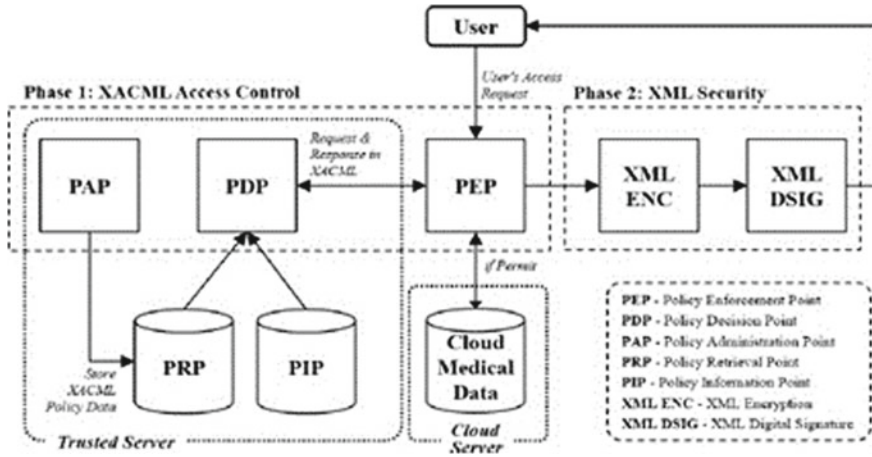
Fig. 4 Accessing records based on XML encryption

3.2 Privacy-Preserving Attribute-Based Mechanism

In this paper, the authors discussed the extensible access control markup language used in a cloud-based EHR model for performing the attribute-based access control which focuses on performing partial encryption, security, and using electronic signatures when a patient's document is sent to a document requester. XML encryption and XML digital signature technology were utilized, as demonstrated in Fig. 4. This method is effective because it sends only the information which is required to the requesters who are authorized to treat the patient.

This model provides the information to the authorized persons who have access to medical records without jeopardizing the confidentiality of patients. The proposed model is categorized into two stages. During the first stage, the access control is performed by using the XACML language. It determines if the user has been granted permission to obtain the medical document. If permission is granted for the user to access the patient's records after access control has been completed. In the second stage, it protects the patient's privacy by sending the documents which are requested by the user in partial encryption and digital signatures.

3.3 Fuzzy-Based Keyword Search in Document

In this method, the query has been provided over the encrypted data with the help of keywords. To achieve this, the data owner needs to encode the keywords and upload the ciphertexts to the cloud server. By providing a trapdoor of a keyword users can able to get the data that is used to search in the encrypted index in the cloud server [11]. There are two methods of searchable encryption, namely, Symmetric

Searchable Encryption (SSE) and Public Encryption Keyword Search (PEKS) [12]. Keyword search in a multi-user setup is a more prevalent case in a cloud environment, where the files are shared by a data owner with a group of authorized users. The persons who have access rights are able to create legitimate trapdoors and search the shared files for keywords. Through keyword search, a fine-grained access control method is created with the help of the ABE method. Through keyword search, ABE is used to provide fine-grained access control. The concept of "multi-key searchable encryption" (MKSE) is proposed as a conceivable approach for solving the scalability issue.

To search a trapdoors keyword in the encrypted document with different keys the MKSE provides a single keyword trapdoor to the user for search over in the server. By using a different keys search with a keyword trapdoor, it will not increase the number of trapdoors that relate to the number of documents. To overcome this issue the edit distance in the encrypted keywords has been presented as a new technique called fuzzy keyword search over encrypted data is identified. But, it cannot address how the structure would work in a multi-user system with changes in searching privileges.

3.4 Dynamic Multi-keyword-Based Search in Cloud

The storing and searching of data in the electronic health record is the most important criteria to perform the task that needs to do. As shown in Fig. 5, a secure way in this Cloud provides adequate storage facilities for enterprises to move their precious information/documents to the Cloud. So, the user can transfer the valued data to the cloud. But, in this transformation security is a major concern between the storage service provider and the user.

In this, cryptographic technique is used for the searchable encryption-based document because it ensures security. The dynamic multi-keyword-based search algorithm, which includes modified completely homomorphic encryption and prims Algorithm, is proposed here. In this, keywords are sorted and a searchable index
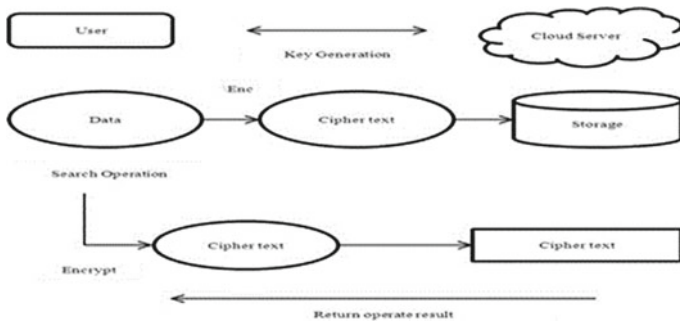
Fig. 5 Cloud data security scheme

for keywords and files is produced based on the keyword frequency contained in the document.

To construct the dynamic tree based on keywords and a file Prim's approach is used. By using a two-way hash table, the encrypted files in the tree are indexed which provides efficient search and document retrieval in a better way. Modified rring-basedcompletely homomorphic encryption is used to encrypt the files. This approach delivers 80 to 86 percent throughput for encrypted and decrypted files, respectively, with indexing time in less for indexing the files than existing approaches [6].

4 Conclusion

In this, we investigated the epidemiological features and preventative approaches of patients with an asymptomatic infection of COVID-19 are identified. However, there is a paucity of scientific evidence, and the specific features of silent infections need to be clarified. Then, we discussed the importance of the design and development of a novel EHR system that should be implanted based on cloud technology for physicians with easy access to records of patient history with the various security mechanism. As part of our review, works need to be taken with the design to provide an interactive tool in a secure manner to be taken.

References

1. Gao Z et al (2020) A systematic review of asymptomatic infections with COVID-19. J Microbiol Immunol Infect
2. Tang N et al (2020) Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy. International Society on Thrombosis and Haemostasis, Wiley
3. Ullah W MD et al (2020) COVID-19 complicated by acute pulmonary embolism and right-sided heart failure. Elsevier, April 2020.
4. Samuel V et al (2019) Design and development of a cloud-based electronic medical records (EMR) system, data, engineering and applications, pp 25–31
5. Zhao et al (2020) A five-compartment model of age-specific transmissibility of SARS-CoV-2. Infect Dis Poverty 9:117
6. van der Hoek L et al (2004) Identification of a new human coronavirus. Nat Med 10(4)
7. Moorthy V et al (2020) Data sharing for novel coronavirus (COVID-19): Bull World Health Organ
8. Goyal V, Pandey O, Sahai A, Waters B (2006) Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM conference on computer and communications security. ACM, pp 89–98
9. Ahmadi M, Aslani N (2018) Capabilities and advantages of cloud computing in the implementation of electronic health record. Acta Inform Med 26(1):24–28
10. Xhafa F et al (2014) Designing cloud-based electronic health record system with attribute-based encryption. Multim Tools Appl

11. Liu Z et al (2015) Cloud-based electronic health record system supporting fuzzy keyword search. Springer Methodologies and Applications
12. Esposito C et al (2014) Smart cloud storage service selection based on fuzzy logic, theory of evidence and game theory. IEEE Trans Comput
13. Seol K et al (2018) Privacy-preserving attribute-based access control model for XML-based electronic health record system. IEEE translations and content mining, vol 6
14. Cao N, Wang C, Li M, Ren K, Lou W (2014) Privacy-preserving multi- keyword ranked search over encrypted cloud data. IEEE Trans Parallel Distrib Syst 25(1):222–233
15. Palanivel Rajan D et al (2017) Dynamic multi-keyword-based search algorithm using modified based fully homomorphic encryption and Prim's algorithm. Cloud Comput

Contourlet Transformed Image Fusion Based on Focused Pixels

M. Ravi Kishore, K. Madhuri, D. V. Sravanthi, D. Hemanth Kumar Reddy, C. Sudheer, and G. Susmitha

Abstract In these works, multi-focus image fusion is performed using the contourlet transform and the removal of residual image components. The main goal of this type of fusion is to integrate the focused pixels and combine them into a fused output image. The most difficult aspect of this method is removing the pixels that are not focused. This methodology is implemented using the MATLAB computation tool, which is included with the image acquisition tool box and the digital image tool box. The experiments conducted here are primarily concerned with image fusion and pixel analysis.

Keywords Image fusion · Counterlet transforms · Focused pixel

1 Introduction

An accurate and reliable composite image is created by combining images from multiple sources. Using the multi-focus picture fusion approach, you can construct a completely focused image from the source photographs when you cannot get all the objects in focus or when the images are only partially focused. Medical imaging, microscopic imaging, remote sensing, computer vision, and robotics all benefit from multi-focus image fusion. The benefits of picture fusion include enhanced spatial information, greater target detection and recognition accuracy, reduced workload, and increased system reliability. Multi-focus combined images are therefore better suited for tasks like segmentation, feature extraction, and object recognition on computers and for human visual perception. Research on information fusion technologies costs the United States roughly $1 billion a year. The study of image fusion is therefore an important one in the field of image processing all around the globe. In geoscience and remote sensing, image fusion applications abound, where satellite pictures of multiple bands and resolutions are combined to derive more relevant information

M. Ravi Kishore (✉) · K. Madhuri · D. V. Sravanthi · D. Hemanth Kumar Reddy · C. Sudheer · G. Susmitha
Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: mrks@aitsrajampet.ac.in

from the Earth's surface. One of the most important defence-related uses of image fusion is the identification of changes over time in images. Due to the increasing growth of medical research, an increase in the demand for imaging research for medical diagnosis, and the availability of multi-modal medical imagery for medical applications, image integration has become easier than ever before. This has resulted in medical image fusion becoming an increasingly popular area of study. CNNSR paradigm is introduced for multi-focus image integration. Local consistency between adjacent patches is taken into account during fusing. Non-negative vocabulary is built using a small number of elements. A patch-level consistency rectification approach is provided during the fusion process in order to remove spatial artefacts. In terms of calculation, the fusion method is quite efficient.

2 Literature Review

As a research area, multi-focus image fusion is characterised by the following: instantaneous measurement of uncertainty; the ability to deal with the number of parameters; and the use of joint probability density functions.

Picture clarity was evaluated using a variety of focus metrics for multi-focus image fusion in this work. A real-time fusion system with quick reaction and robustness can implement all of these spatially defined focus metrics. Focus metrics are evaluated in this research on the basis of their ability to distinguish between image blocks that are focused and those that are defocused. SML (sum-modified-Laplacian) outperforms other focus measures when the execution time is not taken into account, according to experiments on multiple sets of photos.

As a result of numerous successful computer vision and image processing applications, sparse representation (SR) has gained substantial interest in multi-sensor picture fusion. SR learns an over-complete dictionary from a collection of training images for image fusion, resulting in more stable and meaningful representations of the source images than typical MSTs, which assume the basic functions. As a result, the SR-based image fusion approaches outperform the standard MST image fusion methods in both subjective and objective assessments. These images are also easier to use because they are less prone to mis-registration between the source photos. Multi-sensor image fusion using SR-based methods is examined in this review study, which highlights the advantages and disadvantages of each methodology. Algorithmically, we begin by examining the system as a whole from three key points of view: firstly, models of sparse representation; secondly, methods of dictionary learning; and finally, activity levels and fusion rules. That was just the beginning of our discussion, and we will show you how other researchers have tackled similar issues and come up with the right fusion rules for each application, such as multi-focus picture fusion and multi-modality image fusion. Lastly, we run a series of experiments to see how these three algorithmic components affect fusion performance in different contexts. If you are interested in learning more about sparse representation theory, this article is a good starting point.

3 Methodology

3.1 Existing Methods

This approach of merging multi-focus images, known as MST-based methods, is remarkably close to how the human visual system works; hence, the combined images it produces usually have good visual performance, such as crisp and intact focus boundaries. A number of studies have found that NSCT's flexible multiscale, multi-direction, and translation invariance make it an ideal fusion tool. Some disadvantages of the NSCT-based approaches include the unavoidable loss of useful information during picture decomposition and reconstruction, leading to incomplete or inaccurate information in fusion results. This is the most significant limitation. The focus information may be efficiently detected using the focus detection methodology, a popular SPD-based fusion technique. Each of the source photos is first focused on, and then, the decision map is used to build the fusion results. As a focus detection approach, the focus area is detected quite thoroughly. To be fair, as compared to MST-based methods, the focused boundaries produced by these techniques do not have the visual completeness and authenticity of genuine borders. Artefacts and blurring at the boundaries of the fused pictures can be produced as a result of these variations.

"Multi-Focus Image Fusion Based on Nonsubsampled Contour Let" is a new way to overcome the constraints of the previously listed difficulties. Infrared and visible pictures are used in this technique. Infrared (IR) imaging technology is used to determine the temperature of an object. To our eyes, everything emits infrared radiation, which is predominantly emitted by electromagnetic radiation. It is possible to feel heat via infrared radiation (IR). IR radiation is a measure of how hot a thing is.

Clouds, aerosols, atmospheric gases, and the Earth's surface are shown in visible images to show how much sunlight they reflect back into space. Thicker clouds have a higher reflectivity (or albedo) and look brighter than thinner clouds on a visible image.

A digital image's contrast and brightness can be tweaked with the use of a histogram. Finding the fusion model's maximum probability region is an important first step in the detection process. The histogram of the infrared and visible images yields two unique peaks, one for the infrared or IR image and the other for the visible images, which can be used to estimate the maximum probability zone.

3.2 Block Diagram

Image processing techniques like "pixel-hierarchy feature and histogram analysis" are depicted in the block diagram of Fig. 1. The original photos are subjected to band-pass and low-pass filters. Make the photograph look more polished by using these filters. In order to preserve low-frequency information, the low-pass filter reduces the

amount of high-frequency information. A high-pass filter, on the other hand, reduces all low-pass frequency components while keeping high-pass frequency elements.

Using a series of filters, an initial fused image and residual regions are generated; by subtracting these two images, we get the final fused image. These give a clear picture of what is going on. The pixel values of each image are different.

Since they compliment each other to some extent, NSCT and focus detection techniques can be used together for the best results. The fusion of NSCT and focus detection approaches can take advantage of each other's capabilities, resulting in high-quality outcomes. A unique residual removal method proposed in this paper efficiently keeps the focused information of source images in the fused output and greatly eliminates the border problem. A hybrid strategy or an upgraded SPD-based system might be deemed our method. Effectively avoiding the appearance of residuals, this method provides a real focus. The genuine image is known as the conjugate image.

We next do picture histogram analysis and pixel-level analysis on both visible and infrared images in order to prepare the fused output that we need for our analysis.

Algorithm for Grey Scale Image

1. Image with multiple focuses' input: Source images (f1 and f2).
   Fused image as output;
2. Initialisation of two source images, f1 and f2.
3. Show the source images that have been converted.
4. Using 'If' condition to determine the size.
5. If the size is greater than three, convert images from RGB to greyscale.
6. Using the NSCT fusion algorithm to fuse and convert greyscale images.
7. Removing blur from images using Gaussian filters.
8. At the end, the fused image is obtained.
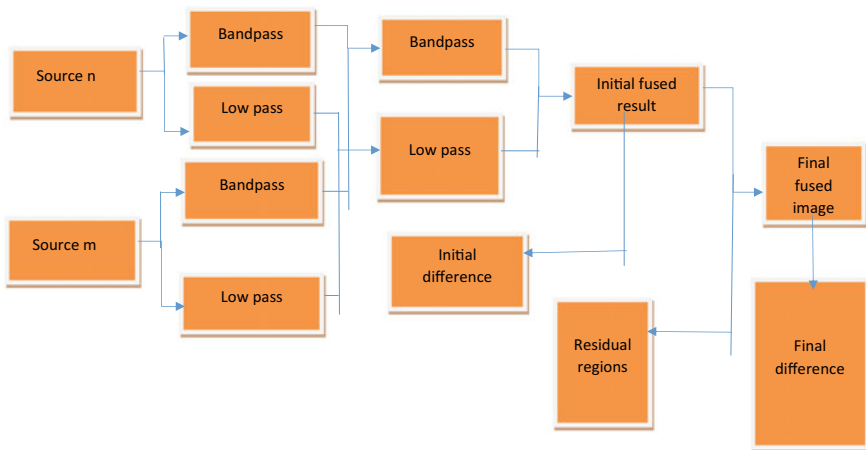
Fig. 1 Multi-focus image fusion

Algorithm for RGB Image

1. Starting two source files (images f1, f2).
2. Image conversion to double precision.
3. Show the source images that have been converted.
4. Iteration is used to obtain the fused colour images.
5. Using the Gaussian filter to remove blur from images.

3.3 Implementation Tool

MATLAB is used in this project to perform image processing computations. MATLAB might be an artificial language. Whenever issues and answers are presented through natural numerical documentation, this software integrates computation, drawing, and associate degreed programming.

By the requirement for comprehensive testing, computerised image preparation shows a range of possibilities for solving problems. Photograph preparation frameworks are known for their extensive testing and experimentation that is likely required before an acceptable relationship is found. Fast mannequin optimistic configurations often play an important function in minimising the rate and time necessary to get at compatible framework execution, as implied by this trademark.

4 Experimental Results

Figure 2a depicts a source image that, when subjected to bandpass and low-pass filters, undergoes a specific process; the bandpass filter is a combination of high pass and low pass. It only allows a limited range of frequencies to pass through. The low-pass filter attenuates high-frequency components while preserving low-frequency components. The initial fused result is obtained, and the residual regions are subtracted before it is given to the final fused image, from which we can find the difference. Figure 2b shows the pixel values of the source image.

Another source image is shown in Fig. 3a. The output will be formed by combining these two images. Figure 3b is the pixel value for Fig. 3a. The flag is clearly visible here, but the observer appears as a blurred image.

By integrating the two photos (1a and 2a) above, we can get the final image as shown in Fig. 4a and their relative pixel values as shown in Fig. 4b. The same procedure can be performed to black-and-white images, as demonstrated in the example below.

In black-and-white photos, the same procedure will occur. As seen in the diagram above, each image has a set of pixel values. The source image is shown in Fig. 5a, and the pixel values are shown in Fig. 5b. Figure 6a is also a source image. Combination of Figs. 5a and 6a will form the overall final output. We can tell the difference between the source image and the final output image by looking at the diagram above.
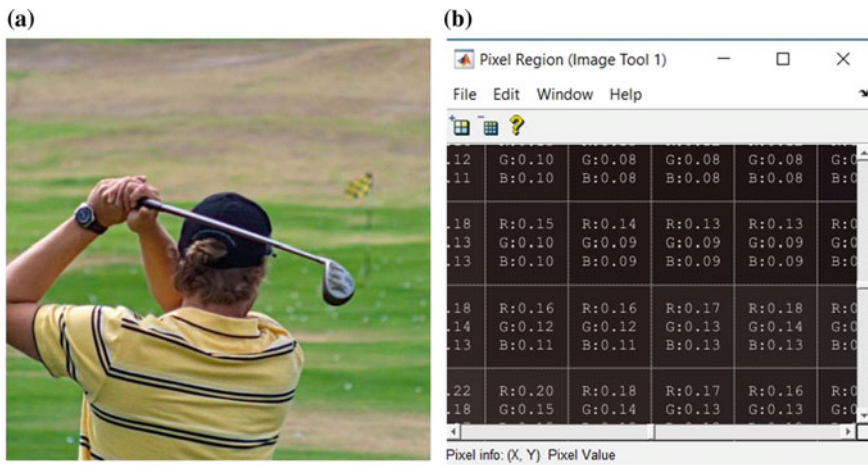
Fig. 2 **a** Source image, **b** pixel value
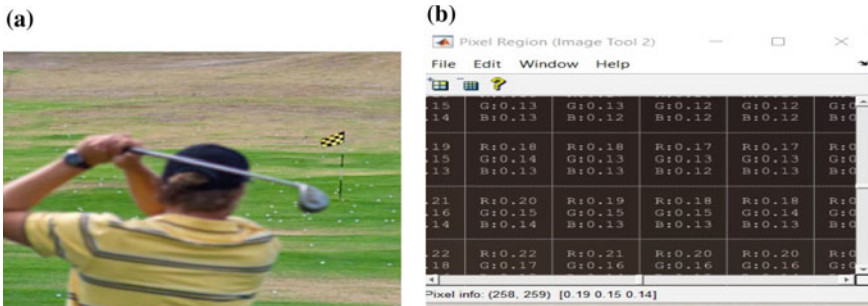
Fig. 3 **a** Source image, **b** pixel value
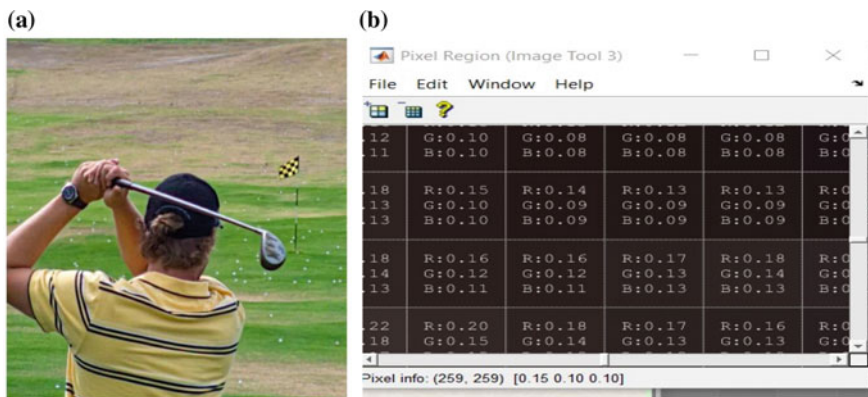
Fig. 4 **a** Final image, **b** pixel value
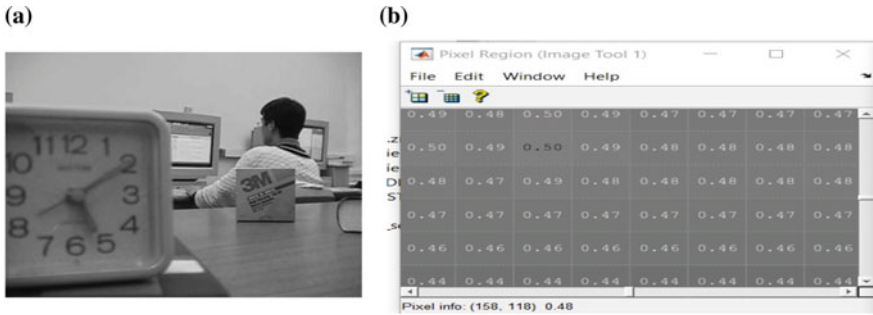
Fig. 5 **a** Source image, **b** pixel value
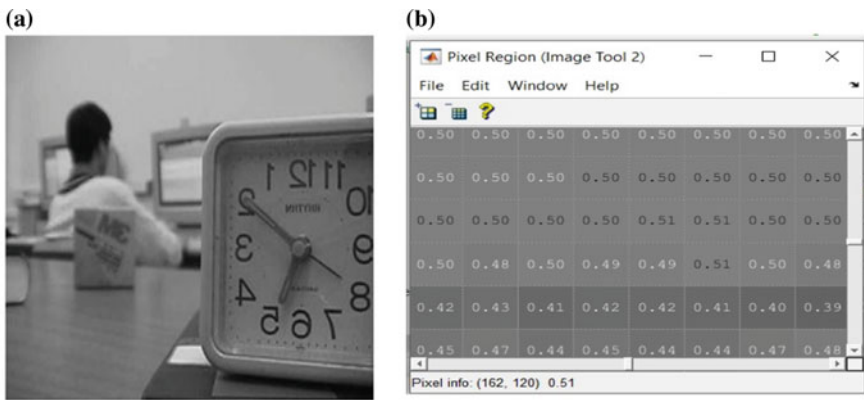
Fig. 6 **a** Source image, **b** pixel value

Observing the aforementioned figures, pixel values can be plainly distinguished (Fig. 7).

5 Conclusion

In this paper, we describe an image fusion approach based on the contourlet transform and image residual component elimination. The incorrectly focused images are erased using the described approach. The MATLAB calculation tool is included in both the image capture tool box and the image processing tool box, and it is via the use of these programmes that the final image is formed. The suggested method's main goal is to accomplish image fusion and pixel values. In comparison to the current method, our method produces a greater number of experimental results. Using the proposed method, we can clearly see both the source photos and the final fused image.
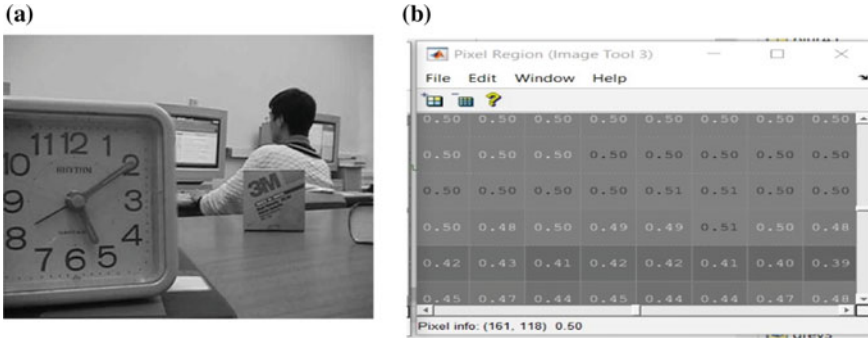
Fig. 7 **a** Final image, **b** pixel value

Bibliography

1. Li H, Wang Y, Yang Z, Wang R, Li X, Tao D (2020) Discriminative lexicon learning-based multiple element decomposition for detail-preserving rackety image fusion. IEEE Trans Instrum Meas 69(4):1082–1102
2. Zhang Z, Blum R (1999) A classification of multiscale decomposition-based image fusion schemes with a performance evaluation for a camera application. Proc. IEEE 87(8):1315–1326
3. Li H, Yang M, Yu Z (2021) Joint image fusion and super-resolution for improved visualisation using semi-coupled discriminative lexicon learning and advantage embedding. Neurocomputing 422:62–84
4. Li H, He X, Yu Z, Luo J (2020) Noise-resistant image fusion with low-rank thin decomposition radio-controlled by external patch previous. Inf Sci 523:14–37
5. Zhang Y, Yang M, Li N, Yu Z (2020) Analysis-synthesis lexicon attempt learning and patch prominence live for image fusion. Signal Method 167:1–13
6. Guo L, Cao X, Liu L (2020) Dual-tree biquaternion rippling rework and its application to paint image fusion. Signal Method (171):107513
7. Liu Y, Wang L, Cheng J (2020) Multi-focus image fusion: a state-of-the-art survey. Inf. Fusion 64:71–91
8. Yu B, Jia B, Ding L, Cai Z, Wu Q, Law R et al (2016) Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion. Neurocomputing 182:1–9
9. Do MN, Vetterli M (2005) IEEE Trans Image Process 14(12):2091–2106
10. Liu S, Jie W, Lu Y, Li H, Zhao J, Zhu Z (2019) Multi-focus image fusion using an adaptive dual-channel spiking cortical model in a non-subsampled shearlet domain. IEEE Access 7:56367–56388

Design of Efficient High-Speed Low-Power Consumption VLSI Architecture for Three-Operand Binary Adders

K. Naveen Kumar Raju, K. Sudha, G. Sumanth, N. Venu Bhargav Reddy, and B. Sreelakshmi

Abstract The three-operand binary adder is a fundamental component in cryptography and pseudorandom bit generators (PRBGs). The carry-save adder (CS3A) is a popular method for performing three-operand addition. In the final stage of the carry saving adder, the ripple-carry adder is used, resulting in a significant critical path delay. For three-operand addition, other adder like Han-Carlson adder (HCA) can be employed, which greatly reduces the delay, but area will increase. Hence, a new high-speed and area-efficient three-operand binary adder is designed, which uses pre-computation logic and carry prefix calculation logic for designing three-operand binary adder with less time and area. In comparison with the previous adders such as the carry-save adder and the Han-Carlson adder, the proposed adder uses less space and time. Xilinx ISE 14.7 tool is used to verify the synthesis and simulation.

Keywords Carry-save adder · Ripple-carry adder · Han-Carlson adder · Xilinx

1 Introduction

Implementing cryptographic algorithms on hardware [1–3] is required to ensure performance of the system for maintaining security. Various cryptographic methods typically use modular arithmetic for arithmetic operations such as modular exponentiation, modular multiplication, and modular addition [4]. The carry-save adder (CS3A) is a commonly used and area-efficient mechanism for performing three-operand binary addition which is utilized in cryptography algorithms to improve security operations.

By combining two-operand adders, three-operand adders can be created. Parallel prefix adders are used to enhance the speed of the operation. Amongst all type of parallel prefix adder, Han-Carlson is majorly used. When the bit size grows, the Han-Carlson adder (HCA) is the fastest.

K. Naveen Kumar Raju (✉) · K. Sudha · G. Sumanth · N. Venu Bhargav Reddy · B. Sreelakshmi
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: naveen.babji@gmail.com

A parallel prefix adder that computes the addition of three input operands utilizing four-stage structures is proposed as a three-operand adder approach.

Xilinx and VHDL

Around 1984, Gateway Design Automation Inc. created Verilog as a proprietary hardware modelling language. The original language is said to have been created by combining characteristics from the most popular HDL language at the time, named HiLo, as well as standard computer languages like C. Verilog was not standardized at the time, and the language changed in practically every edition released between 1984 and 1990.

The Verilog simulator was first used in 1985 and was further expanded in 1987. Gateway's Verilog simulator was used for the implementation. Verilog-XL was the first major expansion, which introduced a few capabilities and implemented the controversial "XL algorithm," a particularly efficient way for executing gate-level simulation.

It was the late 1990s. Cadence design system opted to buy gateway automation system, whose major product at the time was a thin film process simulator. Cadence now owns the Verilog language, as well as other Gateway products, and continues to promote Verilog works as both simulator and language.

At the time, Synopsys was promoting the Verilog which is based top-down design technique. Cadence realized in 1990 that if Verilog remained a closed language, standardization forces would eventually force the industry to switch to VHDL. As a result, Cadence established the Open Verilog International (OVI) and gave it the Verilog Hardware Description Language documentation in 1991. The incident that "opened" the language was this one.

- Hardware description language is abbreviated as HDL. A register transfer level (RTL) can be used to represent any digital system, and HDLs are used to define the RTL.
- Verilog is a general-purpose language which is simple to learn and use. This Verilog is similar to syntax of C.
- The goal is to specify how data moves between registers and how the architecture handles it.
- Hierarchical design concepts play an important part in defining RTL. With numerous levels of abstraction, hierarchical design approach improves the digital design flow.
- Verilog HDL may use these layers of abstraction to represent any digital design's RTL description in a simpler and more efficient manner.
- For example, HDL is used to design both switch level and gate-level designs. Switch level is used to explain the layout of resistors, wires, and transistors on an integrated circuit (IC) chip, and gate level is used to design of flip flops and logical gates in a digital system. All these level are implemented with Verilog.
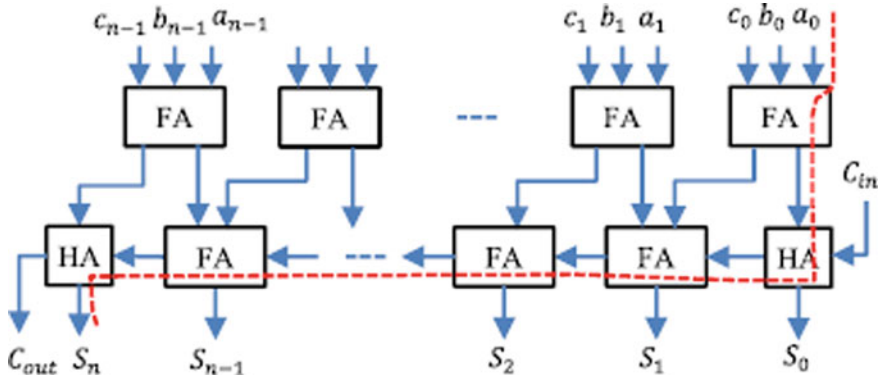
Fig. 1 Carry-save adder (CS3A) showing critical path delay

2 Existing Methods

2.1 Carry-Save Adder

Carry-save adder (CSA) is used for the addition of three operands [9, 10, 11–14]. It does a two-stage addition of three operands. The first stage consists of full adders. From the three binary inputs a_i, b_i, and c_i, each full adder generates the carry and sum bit. The ripple-carry adder is the second stage. At the output of three-operand addition, ripple-carry adder generates the one-bit carry out and total sum signals. In the ripple-carry stage, the "carryout" signal is propagated across the n number adders as shown in Fig. 1. When a result, as the bit length grows, the delay increases linearly. A dashed line indicates where the crucial route delay occurs. It demonstrates that critical path delay is determined by the ripple-carry stage's carry propagation delay.

2.2 Han-Carlson Adder

The Han-Carlson adder features two Brent-Kung stages at the start and conclusion, as well as Kogge-Stone stages in the centre. This Han-Carlson adder has small delay compare to other methods but it has more hardware complexity.

The addition of three operands is computed in two steps using two-operand Han-Carlson adders (HC2A-1 and HC2A-2) as shown in Fig. 2. Base logic, sum logic, and propagate and generate (PG) logic are the three layers of the Han-Carlson adder.

The downside is that the adder's area grows as the bit length grows. Hence, a new high-speed area-efficient three-operand binary adder technology and the VLSI architecture are designed to reduce this trade-off between area and latency.

Fig. 2 Han-Carlson adder (HCA)
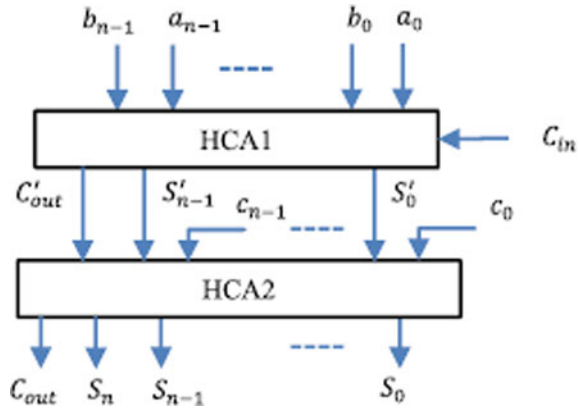
3 Proposed Method

The new proposed three-operand adder is described in this section. It makes use of a parallel prefix adder. In order to compute the addition of three operands, it has four steps rather than three. Logics for bit-addition, base, propagate and generate (PG), and sum are the four steps in the process. Each of these four steps is expressed logically as follows:

Step I: Logic for bit addition

$$S'_i = a_i \oplus b_i \oplus c_i,$$
$$cy_i = a_i \cdot b_i + b_i \cdot c_i + c_i \cdot a_i$$

Step II: Logic for base

$$G_{i:i} = G_i = S'_i \cdot cy_{i-1}, \quad G_{0:0} = G_0 = S'_0 \cdot C_{\text{in}}$$
$$P_{i:i} = P_i = S'_i \oplus cy_{i-1}, \quad P_{0:0} = P_0 = S'_0 \oplus C_{\text{in}}$$

Step III: Logic for propagate and generate (PG)

$$G_{i:j} = G_{i:k} + P_{i:k} \cdot G_{k-1:j},$$
$$P_{i:j} = P_{i:k} \cdot P_{k-1:j}$$

Step IV: Logic for sum

$$S_i = (P_i \oplus G_{i-1:0}), \quad S_0 = P_0, \quad C_{\text{out}} = G_{n:0}$$

The three-operand binary adder's proposed VLSI architecture and internal organization are depicted in the diagram below. There are four phases to the three-operand

binary addition. The first level is bit-addition logic, which is made up of an array of complete adders, each of which calculates the sum (S_i) and carries (cy_i) signals as shown in Fig. 3.

The first stage full adder's output signal sum (S_i) bit and carry (C_{in}) bit are given as input to the second stage that is base logic used for computation of generate (G_i) and propagate (P_i) signals as shown in Fig. 4.

$$G_{i:i} = G_i = S_i' \cdot cy_{i-1};$$
$$P_{i:i} = P_i = S_i' \oplus cy_{i-1}$$

The third stage in the proposed adder is propagate and generate (PG) logic, which is addition of grey and black cells that computes the carry bit as shown in Fig. 5. The logical expression of propagate and generate is shown below:
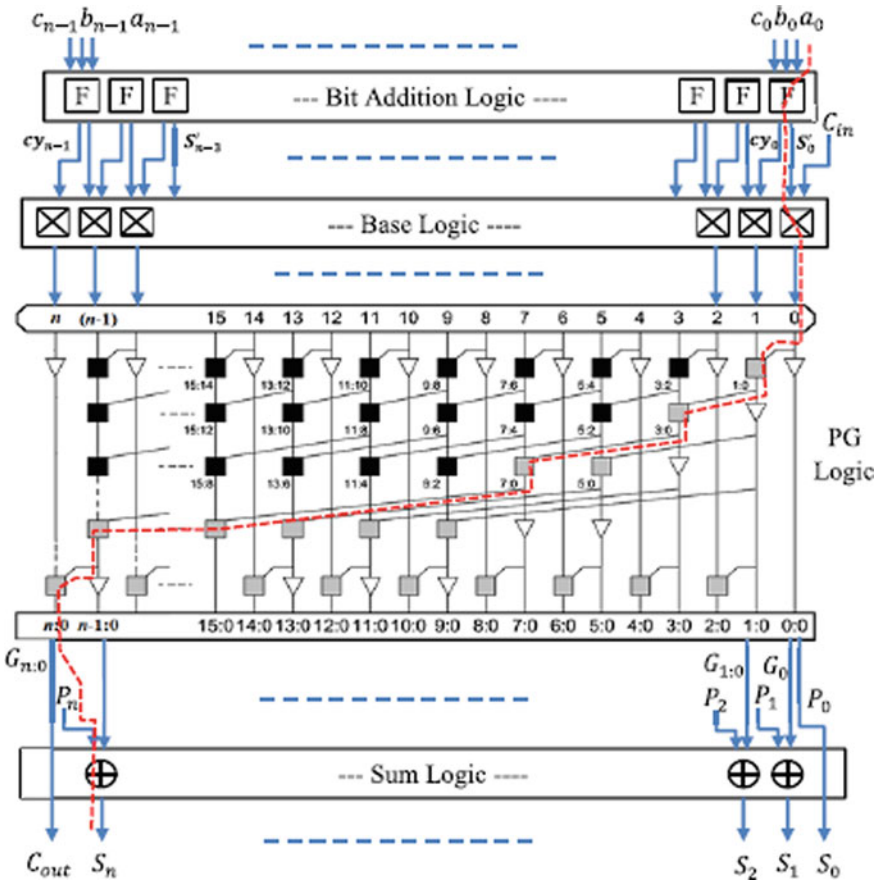
Fig. 3 Suggested three-operand binary adder
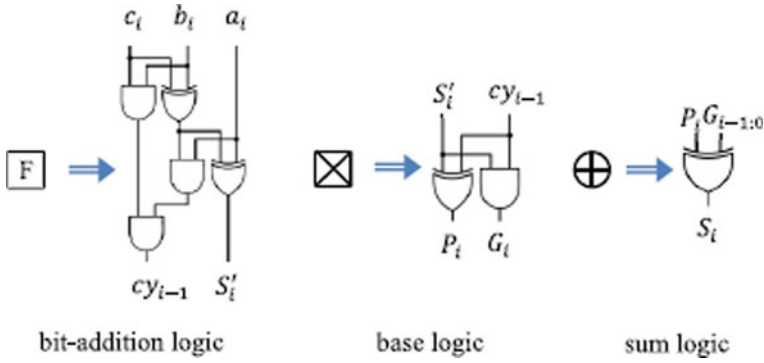
Fig. 4 Logic diagram for bit addition, base, sum logics
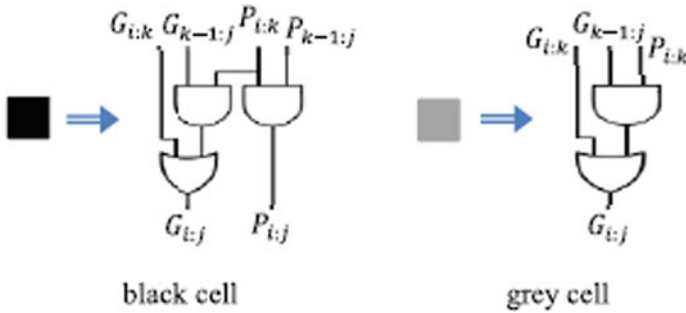
Fig. 5 Logic diagram black-cell and grey-cell

$$G_{i:j} = G_{i:k} + P_{i:k} \cdot G_{k-1:j},$$
$$P_{i:j} = P_{i:k} \cdot P_{k-1:j}$$

The proposed adder has $(\log 2\ n + 1)$ prefix computation steps. The final stage in the proposed adder is sum logic which computes "sum (S_i)" by using the logical expression

$$S_i = (P_i\ G_i - 1 : 0)$$

Here, propagate (P_i) and generate (G_i) are the outputs of PG logic, and carryout signal is generated from the carry generate bit G_n:0.

The simulation is initiated by generating the block diagram as shown in Fig. 6 where the corresponding pins with respect to the order values and also sum output are represented, and the proposed adder technology schematic is provided by Fig. 7.

Fig. 6 Block diagram
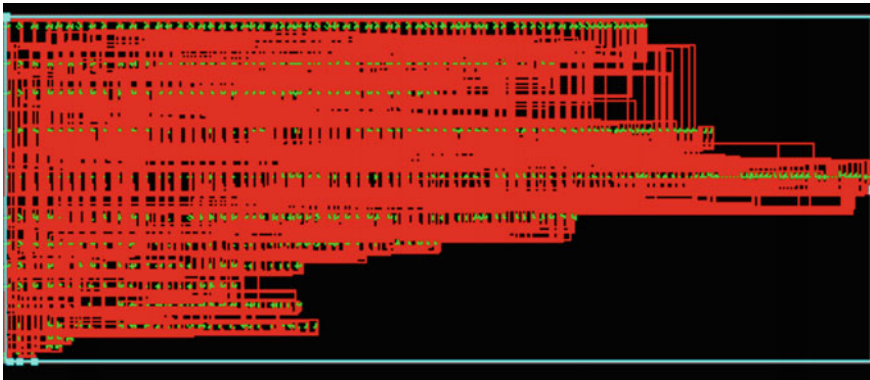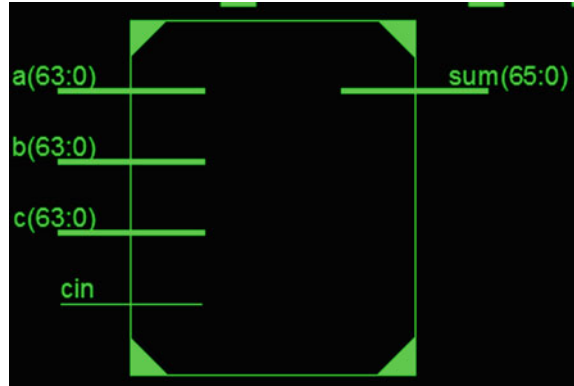
Fig. 7 Proposed adder technology schematic

4 Simulation Results

See Figs. 8, 9 and 10 and Table 1.

5 Conclusion

The addition of three operands is used in several cryptography algorithms. Hence, area-efficient and high-speed three-operand binary adder is required. The suggested three-operand adder is a parallel prefix adder that uses four steps for the addition of three operands utilizing pre-computation logic. The proposed adder reduces delay and area. In comparison with the existing methods three operand adder like CS3A and HC3A adders, the proposed three-operand binary adder has small delay and area. The simulation results are verified with the help of Xilinx ISE tool.

Fig. 8 Proposed three-operand binary adder simulation results

```
Slice Logic Utilization:
 Number of Slice LUTs:                      368  out of  204000      0%
    Number used as Logic:                   368  out of  204000      0%

Slice Logic Distribution:
 Number of LUT Flip Flop pairs used:        368
    Number with an unused Flip Flop:        368  out of     368    100%
    Number with an unused LUT:                0  out of     368      0%
    Number of fully used LUT-FF pairs:        0  out of     368      0%
    Number of unique control sets:            0

IO Utilization:
 Number of IOs:                             259
 Number of bonded IOBs:                     259  out of     600     43%
```

Fig. 9 Screenshot of area report

```
Data Path: b<0> to sum<65>
                         Gate     Net
  Cell:in->out   fanout  Delay   Delay  Logical Name (Net Name)
  -------------------------------------  ------------
   IBUF:I->O         3   0.000   0.507  b_0_IBUF (b_0_IBUF)
   LUT3:I0->O        2   0.043   0.608  g1/carry1 (c1<0>)
   LUT5:I0->O        4   0.043   0.630  pg1/gc1/G1 (x<1>)
   LUT6:I0->O        2   0.043   0.618  pg1/gc33/G1 (x<2>)
   LUT6:I0->O        4   0.043   0.512  pg1/gc2/G1 (x<3>)
   LUT6:I3->O        3   0.043   0.534  pg1/gc4/G1 (x<7>)
   LUT6:I2->O        1   0.043   0.522  pg1/gc8/G1 (pg1/gc8/G)
   LUT4:I0->O        3   0.043   0.417  pg1/gc8/G2 (x<15>)
   LUT6:I4->O        1   0.043   0.405  pg1/gc16/G1_SW0 (N4)
   LUT6:I4->O        5   0.043   0.428  pg1/gc16/G1 (x<23>)
   LUT6:I4->O        1   0.043   0.405  pg1/gc28/G111 (pg1/gc28/G11)
   LUT5:I3->O        5   0.043   0.518  pg1/gc28/G112 (x<31>)
   LUT6:I3->O        2   0.043   0.527  pg1/gc32/G4 (pg1/gc32/G3)
   LUT6:I2->O        1   0.043   0.339  h1/carry1 (sum_65_OBUF)
   OBUF:I->O             0.000          sum_65_OBUF (sum<65>)
  -------------------------------------
   Total                         7.531ns (0.559ns logic, 6.972ns route)
                                 (7.4% logic, 92.6% route)
```

Fig. 10 Screenshot of delay report

Table 1 Comparison between existing methods and proposed method

	Area (in LUTs)	Delay (in ns)
CS3A	192	36.786
HC3A	444	9.622
Proposed adder	368	7.531

References

1. Islam MM, Hossain MS, Hasan MK, Shahjalal M, Jang YM (2019) FPGA implementation of high-speed area-efficient processor for elliptic curve point multiplication over prime field. IEEE Access 7:178811–178826
2. Liu Z, GroBschadl J, Hu Z, Jarvinen K, Wang H, Verbauwhede I (2017) Elliptic curve cryptography with efficiently computable endomorphisms and its hardware implementations for the Internet of Things. IEEE Trans Comput 66(5):773–785
3. Liu Z, Liu D, Zou X (2017) An efficient and flexible hardware implementation of the dual-field elliptic curve cryptographic processor. IEEE Trans Ind Electron 64(3):2353–2362
4. Parhami B (2000) Computer arithmetic: algorithms and hardware design. Oxford University Press, New York
5. Montgomery PL (1985) Modular multiplication without trial division. Math Comput 44(170):519–521
6. Kuang S-R, Wu K-Y, Lu R-Y (2016) Very Large Scale Integr (VLSI) Syst 24(2):434–443
7. Kuang S-R, Wang J-P, Chang K-C, Hsu H-W (2013) Energy-efficient high-throughput montgomery modular multipliers for RSA cryptosystems. IEEE Trans Very Large Scale Integr (VLSI) Syst 21(11):1999–2009
8. Erdem SS, Yanik T, Celebi A (2017) A general digit-serial architecture for montgomery modular multiplication. IEEE Trans Very Large Scale Integr (VLSI) Syst 25(5):1658–1668
9. Katti RS, Srinivasan SK (2009) Efficient hardware implementation of a new pseudo-random bit sequence generator. In: Proceedings of the IEEE International Symposium on Circuits and System, Taipei, Taiwan, pp 1393–1396
10. Panda AK, Ray KC (2019) Modified dual-CLCG method and its VLSI architecture for pseudorandom bit generation. IEEE Trans Circuits Syst I, Reg Papers 66(3):989–1002

DLCNN Model with Multi-exposure Fusion for Underwater Image Enhancement

Biroju Papachary, N. L. Aravinda, and A. Srinivasula Reddy

Abstract Oceans comprise a huge portion of our planet's surface, and these water resources determine our planet's health. Underwater flora and fauna research is an essential component of oceanographic research. In order to create a more aesthetically attractive image, underwater image improvement is a crucial element. There are a variety of image processing techniques available, most of which are extremely easy and quick to use. However, traditional methods failed to provide the highest level of precision. Due to haze caused by light reflected from the surface and separated by the water particles, the underwater picture captured by the camera has low prominence, and colour variation is caused by attenuation of different wavelengths. To address these shortcomings, the proposed study employs the advanced deep learning convolution neural network (DLCNN) model. The network is trained and evaluated using features based on the Laplacian and Gaussian pyramids. These characteristics are used to change the parameters of the multiple exposure fusion (MEF), allowing for precise improvement. The proposed DLCNN-MEF technique gives better results when compared with state-of-the-art approaches technique with respect to the various parameters including such as PSNR, BRISQUE and MSE values for comparing the results.

Keywords Contrast limited adaptive histogram equalization · Ycbcr colour space · High resolution · Deep learning convolution neural network

1 Introduction

Underwater images get degraded due to poor lighting conditions [1] and natural effects like bending of light, denser medium, reflection of light and scattering of light, etc. As the density of sea water is 800 times denser than air, when light enters from air

B. Papachary (✉) · N. L. Aravinda · A. Srinivasula Reddy
ECE Department, CMR Engineering College, Medchal, Hyderabad, India
e-mail: biroju.chary@gmail.com

N. L. Aravinda
e-mail: aravindanl@cmrec.ac.in

179

(here lighting source) to water, it partly enters the water and partly reflected reverse. More than this, as light goes deeper in the sea, the amount of light enters the water also starts reducing. Due to absorption of light in the water molecules, underwater images will be darker and darker as the deepness increases. Depending on the wavelength also, there will be colour reduction. Red colour attenuates first followed by orange. As the blue colour is having shortest wavelength, it travels longest in seawater there by dominating blue colour to the underwater images affecting the original colour of the image. The overall performance of underwater [2] vision system is influenced by the basic physics of propagation of light in the sea. The overall performance of underwater vision system is influenced by the basic physics of propagation of light in the sea.

One must take care about absorption [3] and scattering [4] of light when considering underwater optical imaging systems. Most often, accurate values of attenuation and thorough knowledge of forward and backward scatter are required for restoring underwater images performance of underwater vision system which is influenced by the basic physics of propagation of light in the sea. One must take care about absorption and scattering of light when considering underwater optical imaging systems. Most often, accurate values of attenuation and thorough knowledge of forward and backward scatter are required for restoring underwater images. Due to apprehension [3] regarding the present conditions of the world's oceans, many large-scale scientific projects have instigated to examine this.

The underwater video sequences are used to monitor marine species [4]. Underwater image is inundated by reduced visibility conditions making images deprived of colour variation and contrast. Due to this restriction, other methods like sonar ranging have been preferred previously. Since alternate methods yield poor resolution images which are hard to understand, nowadays for close range studies, visible imaging is preferred by scientists.

The emphasis of these works deceits primarily in the area of implementation of vision system [5] which involves analysis of enhanced images. Image enhancement techniques [6] are usually divided in frequency domain and spatial domain methods. The frequency domain methods [7] are normally based on operations in the frequency transformed image, while spatial domain methods [8] are based on direct manipulation of the pixels in the image itself. To solve the effects, the proposed method is contributed as follows:

1. Laplacian and Gaussian pyramid-based MEF mechanisms are implemented to generate the features for the DLCNN model.
2. DLCNN model is proposed for training and enhancement of underwater images by using the multiple numbers of layers.

Rest of the paper organized as follows, Sect. 2 deals about discussion of various existing methods with their drawbacks. Section 3 deals about detailed operation of proposed mechanism with DLCNN working. Section 4 deals with simulation analysis of proposed method and comparison of results with state-of-the-art approaches. Section 5 concludes the paper with possible future enhancements.

2　Related Works

In the article [9], it is potential to eliminate the complex interference and rebuild the underwater images by enhancement techniques. The underwater images are enhanced by using the algorithms like grey world for clearance and dark channel prior for processing of underwater images by applying the BP network for restoration of details in the underwater image. The TV model procedure is executed because to cover the blank area after the recognition and elimination of object. The article [10] describes the enhancement technique by using the wavelet fusion for underwater images which is due to the absorption and reflection of light while capturing the images in water. By using this implementation, we can evaluate and dehazed or enhance the contrast and colour of the image.

Owing to the scattering and absorption of light colour, alteration is emerged in underwater images. For improving the underwater images, colour distortion CC-Net [11] is used, and for contrast enhancement, we are using HR-Net which consist of single transmission coefficient. These two Nets are the grouping of UIE-Net [12] which is one of the frameworks in CNN architecture. This implementation progresses the learning process and convergence speed simultaneously [13]. It overcomes the several optical transmissions, hazing and colour distortions. For features extraction, these two are trained to the CNN framework. Yadav et al. researches in [14] describe that in order to raise the visibility of the hazy image or video, we utilize the contrast limited adaptive histogram equalization. When compared to other enhancement techniques, MEF algorithm gives better enhancement results because it applies for both colour and grey images for both regions.

A new frame is generated after the adjustment of intensity values over the image is named as histogram equalization. Based on the neighbouring pixel values, AHE [15] adjusts the intensity levels over a particular region of any fog (homogeneous) type of images only. For histogram equalization shape in MEF, distribution parameter is utilized. For noisy image, we are applying clip limit. The light travelling in the water with altered wavelengths leads to the colour variation due to the attenuation. Here, by employing the image processing techniques, we can enhance the underwater images. In this paper, we have both enhancement and restoration operations. There are different types of image enhancement techniques [16] for enhancing the image. Image restoration is nothing but the degradation is used to restore the loss of information. In this paper, UIQM is utilized for measuring the quality of the underwater images. This implementation gives the better-quality results when compared with other enhancement techniques [17].

In [18], authors proposed the deep learning architecture for underwater image enhancement using the white balancing multiple underexposure versions reduction prior (WB-MUV) model. The WB-MUV method for enhancing the underwater images is by using the deep residual architecture. This is one of the networks in CNN which is used to convert the low-resolution image into high-resolution image. The underwater images are less in contrast and colour distortion. Generally, we are considering the input image as reference. First, we consider a cell after training the

data set. Here, we are using the YCbCr colour space for conversion of input image and the whole procedure is performed in Y component only. By utilizing the bicubic interpolation, we can change the image into low-resolution image.

For bicubic interpolation, the block uses the weighted average of four transformed pixel values for each output pixel value. For improving or enhancing the underwater images, there are various filters are used as image enhancement techniques. The network which we are used is a deep learning network. The layers which we are used in a novel architecture are an image input layer, convolution and ReLU layers. Due to this residual network, degradation was occurred and consists of saturated accuracy while using deep residual blocks. These blocks are helpful to transmit the information (data) between layers. By increasing the number of layers, there may be chances to raise the accuracy of the network.

To overcome the drawbacks of the WB-MUV approach, we are using different layers of DLCNN for enhancing the underwater images including such as BN layer, ResNet layer and CycleGAN. The resultant enhanced underwater image has less visibility and contrast. Edge difference loss (EDL) plays a significant role in order to know the pixel difference between the two images. Mostly in case of WB-MUV, usage of MSE loss is more to calculate the peak signal-to-noise ratio (PSNR). The depth which we are using in this process is 20. The low-resolution image is resized based on the reference image size and difference is measured between them. The combination of the resultant upscale and difference images is called as the train data. The residual network is trained after assigning the layers and training options. In order to get the high-resolution image, and the difference image is calculated by considering the test image. For WB-MUV technique of enhancement, there is no need of same size of images are needed. By training the proper data set, it gives the better results of enhanced image.

3 DLCNN Architecture

In order to overcome the outcomes of the related works in this proposed method, a novel architecture is utilized by considering the network depth as 10. This network contains several layers. This proposed method improves more visibility of the enhanced underwater images. The below shows the block diagram of the network which is used for proposed implementation (Fig. 1).

The first layer of this network is image input layer having image data. It applies data normalization by considering this input data into the network. To extract the features from input layer, convolutional layer is used. By taking the image and kernel as inputs, this layer prevents the feature values from neighbourhood pixels by adding a bias to the input of the convolutional layer along with the size of the kernel or filter. Filters are used to improve the operations including such as blur, sharpen and edge detection. Then we are using rectified linear unit (ReLU) layer as activation, where the input value is set as zero which is less than zero by applying the threshold set-up before applying the pooling layer. After applying the convolutional layer, next layer
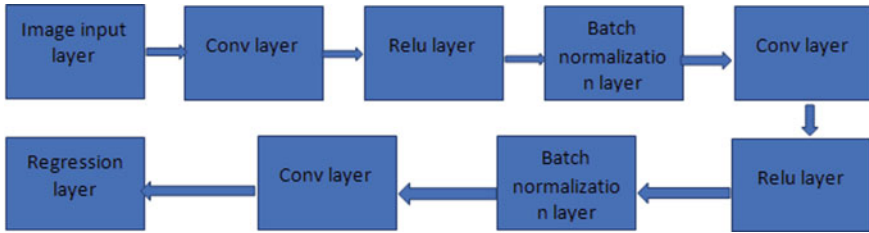
Fig. 1 DLCNN network architecture

will be the batch normalization layer in order to increase the network training process correspondingly by reducing the activations and gradients. The network initialization is applied in order to drop the sensitivity. This layer sustains the network as steadier and it decreases the over fitting problems which is occurs due to the. By calculating the loss and accuracy which are called as activation metrics, we can estimate the over fitting problem. This is main problem in the neural network. And also, we can observe the over fitting whenever the validation loss is more the training loss will be reduced. Then, it calculates the normalized activations as

$$\hat{X}_i = \frac{X_i - \mu_\beta}{\sqrt{\sigma_\beta^2 + \varepsilon}} \tag{1}$$

where X_i is the input and μ_β is the mean and variance σ_β^2.

These two parameters are measured after the training process only. If we consider the pre-trained network, then the mini batch values are applied for mean and variance. In order to consider the mean as zero and variance as one, the activation is computed as

$$y_i = \gamma \hat{X}_i + \beta \tag{2}$$

where γ and β are called as the training parameters.

These two are stored as trained mean and variance values. Finally, we are using the regression layer in this architecture. This layer is used to predict the output from the feature values of the input. To estimate the half, mean square error problems regression layer were utilized. The mean square error for single observation is given as

$$\text{MSE} = \sum_{i=1}^{R} \frac{(t_i - y_i)^2 \beta}{R} \tag{3}$$

In order to calculate the half, mean square error for image to image regression network where H, W, C are the height, width and number of channels and the loss function is given as

$$loss = \frac{1}{2} \sum_{\rho=1}^{HWC} (t_i - y_i)^2 \qquad (4)$$

For regression problems, it will help to increase the network training process. This architecture is easy to implement and gives the more accuracy of training process.

4 Proposed Method

Underwater image contrast enhancement is a method which is used to increase the visibility of images. Here, the intensity of the pixel grey level is modified based on a function. Intensity-based method is given as

$$I_o(x, y) = f(I(x, y)) \qquad (5)$$

The original image is (x, y), the output image is $I_o(x, y)$ and f is the DLCNN transformation function. Intensity-based methods transmute the grey levels over the entire image. Even after the transformation, pixels with same grey levels throughout the image remain same. Contrast stretching is a generally used method that falls into this group. Figure 2 represents the proposed method of underwater image enhancement with contains the white balance-based brightness preserving DLCNN-MEF approach as its operational function $((x, y))$.

The step-wise operation of proposed method is as follows:

Step 1: Underwater images corrupted by lighting conditions may lose the visibility of the scene considerably. Enhancing approach is not available to remove entire haze effects in degraded images. The proposed algorithm includes deriving two inputs from the single degraded image and it recovers colour and visibility of the entire image. Colour correction is applied to the image after the fusion process. In the proposed algorithm, two inputs are derived from a single degraded input. The first derived input is obtained by white balancing the input. This step aims at removing chromatic casts in the image. More attention is given to red channel of the image as it attenuates more in underwater. The second derived input I_2 is capable of enhancing those regions having low contrast. It is obtained by contrast stretching the median filtered input image. Therefore, the first derived input I_1 avoids colour casts and the second derived input improves the uneven contrast. The important features of these derived inputs are computed using different weight maps. The different weight maps derived in the proposed model are exposedness, saliency, Laplacian and colour cast.

Step 2: The first output will be applied as input to the gamma correction using Laplacian pyramid operation. The weight maps play a critical role in the outcome of the final fused result. The weight maps generated should have a non-negative value. The weight maps are some measures of the input image. It represents t
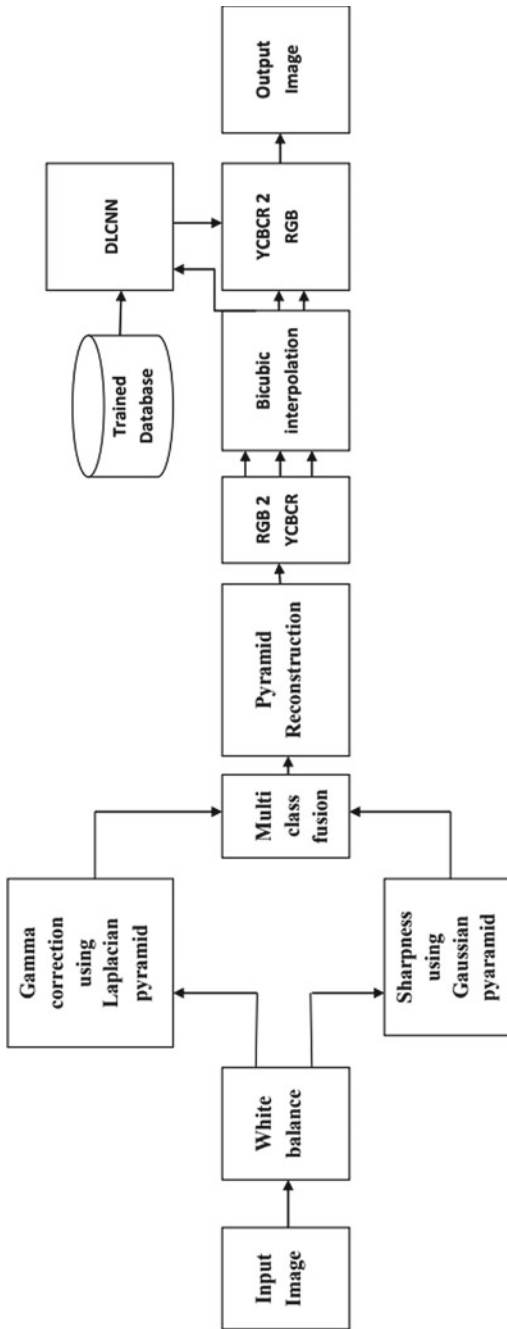
Fig. 2 Proposed DLCNN with MEF block diagram

finer details of the image. The finer details from each image have to be extracted out and fuse them together to form the enhanced image. The weight maps are to be designed carefully to extract the details. New weight maps were used which showed better results in enhancing degraded underwater images. The different weight maps are Laplacian, saliency, colour cast and exposedness.

Step 3: The second input will be applied as input to the sharpening using Gaussian pyramid operation. In an underwater image, all pixels will not be exposed.

Step 4: Multi-scale fusion technique is used to improve the visibility of poorly illuminated areas by using multi-scale fusion-based strategy. The fundamental principle of this method is extracting the best features from the given input image and then to apply the estimated perceptual-based qualities called weight maps to them and finally fuses them together to form the enhanced output. In the fusion-based strategy, two inputs derived from the single image will be by white balancing and by median filtering followed by contrast stretching. White balancing is used to reduce colour casting from the individual input image; estimation of four weight maps is done. One is exposedness weight map, which measures the exposedness of pixels not exposed to underwater constraints. Laplacian weight map assigns high values to textures as well as edges. Colour cast weight map, introduced newly, increases red channel value thereby reducing colour cast. Saliency weight map measures the amount of discernible concerning other pixels. These weight maps are computed and applied to both the derived inputs. This method is a per pixel implementation. Fusion is a process of combining the relevant information from a set of input images into a single image, where the resultant fused image will be more informative and complete than the input images.

Step 5: Pyramid reconstruction is the process of reconstruction of the Gaussian and Laplacian pyramid outcomes and results the pyramid reconstruction output, respectively.

Step 6: The pyramid reconstruction output is applied to RGB2YCBCR operation. It will generate the output as Luminance (Y), chromium blue (CB) and chromium red (CR) outcomes, respectively.

Step 7: Perform the bicubic interpolation operation individually on Y, CB and CR outcomes. These bicubic interpolations will result the perfect region of interest respectively.

Step 8: The detailed operation of DLCNN explained in Sect. 3. The luminance output of the bicubic interpolation is applied to implement mean brightness; we used morphological based operation, which is computationally not exhaustive. We are essentially extracting pixel block centred also referred to as patch at $\Omega(x, y)$. We determine the minimum value for each block. Hence, we get three values corresponding to each colour for every block of pixels. From these three minimum intensity values, we chose the most minimum value and replace it at the centre location of the processed patch $\Omega(x, y)$. This step is repeated till the entire image is operated upon. Finding the minimum value for a pixel block in a grayscale image is same as carrying out a morphological operation. In this case, we can separately apply this operation on individual colour channels corresponding to H, S and I. This step is then followed with finding the minimum out of the three

colour planes for any structuring element. Inspired from mean channel prior, we derive the modified red colour priority depth map image, which is given as

$$I_I^{mean}(x, y) = \min_{\gamma \varepsilon R}\left(\min_{\gamma \varepsilon \cap \Omega(x, y)} (I_R^n(x, y)) \right) \tag{6}$$

where $\Omega(x, y)$ is the structuring element, n corresponds to the number of partitions and I corresponds to the intensity colour channel. This depth map gives the pictorial estimate of the presence of haze in an image and is useful for estimating the mean transmission map.

Step 9: The adaptive histogram equalization is utilized only to remove the noise over the homogeneous areas. In order to overcome this issue, MEF is proposed because it applies each and every pixel of the image. For generation of contrast transform function, there are different types of distributions including such as Rayleigh, uniform and exponential. Rayleigh distribution is the most useful for enhancing the underwater images. This algorithm is only applied for smaller regions or areas throughout the image. To eliminate the boundaries by using the bilinear interpolation the tiles are combined. To prevent the over saturation area in case of homogeneous regions, we are using the contrast factor. The MEF produces the enhanced results in some times without using the clip limit (contrast factor). In this proposed algorithm, we are decreasing the network depth that is 10. And the number of layers which we are used is 10 layers. By decreasing the depth size, the performance of the network is improved and gives the better enhancement results. Finally, DLCNN output is concatenated with other CB and CR components and results the output as the RGB version of contrast enhanced colour image.

5   Results and Discussions

The proposed method is implemented using MATLAB 2018a software simulation on various types of underwater test images such as low light images, dusty environment images, hazy images and general purposed images. The proposed method performs outstanding enhancement and gives the better qualitative and quantitative evaluation compared to the state-of-the-art approaches. The enhancement results of the proposed implementation by using a novel architecture for underwater images is shown and discussed. In this paper, Table 1 is used for improving the contrast and prevents the original data. For effectively calculating the quantitative evaluation, the PSNR, BRISQUE and MSE parameters are calculated as follows:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [In(i, j) - out(i, j)]^2 \tag{7}$$

$$PSNR = 10.\log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{8}$$

$$\text{BRISQUE} = \frac{(2.\varphi_{xy} + c_1)(2.\sigma_{xy} + c_2)}{(\varphi_x^2 + \varphi_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{9}$$

Here, φ_x is the luminance mean of input image, φ_y is the luminance mean of output image and φ_{xy} is combined mean. σ_x is the standard deviation of input image, σ_y is the standard deviation of output image and σ_{xy} is combined standard deviation. Finally, c_1 and c_2 are the contrast comparison functions.

In Table 1, columns represent the various input images, WB-MUV output images and DLCNN-MEF output images and rows represent the input images with resulted output images. By altering the values using the contrast limited adaptive histogram equalization algorithm, the grey scale image contrast will be enhanced. By observing the existing WB-MUV output images, it is clearly observed that the images are affected by radiance effect. This problem is resolved by applying the MEF with DLCNN. The proposed DLCNN-MEF algorithm in this paper is applied to the underwater images which are having low contrast.

Table 2 gives the different parameters values by testing different images. By calculating the parameters like PSNR, BRISQUE and MSE, we compare the results of

Table 1 Performance analysis of proposed method on various test images
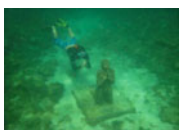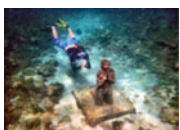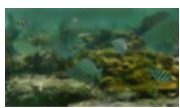
	Input images	WB-MUV [18]	DLCNN-MEF output
Test 1			
Test 2			
Test 3			
Test 4			
Test 5			

Table 2 Parameters comparison between WB-MUV [18] and proposed DLCNN-MEF method

Input	WB-MUV[18]			DLCNN-MEF		
Figures	PSNR	BRISQUE	MSE	PSNR	BRISQUE	MSE
Test 1	17.6907	31.1223	1.23E+03	41.5391	23.183	4.1963
Test 2	16.807	26.6708	1.33E+03	42.0304	31.587	3.0685
Test 3	16.6154	30.0721	1.32E+03	40.7258	31.1171	5.3081
Test 4	16.807	26.6708	1.31E+03	42.8645	40.4827	3.7423
Test 5	15.819	28.037	1.49E+03	41.3099	31.0342	4.132

underwater enhanced images between WB-MUV and proposed implementations. The peak signal-to-noise ratio (PSNR) is defined as the ration between the noise corruption and maximum usable power. If PSNR is more, the enhanced results give better performance. Here, the PSNR is nearer to 42. The errors are calculated by using mean absolute error (MSE). It gives good results for low value of error. Here, the value which is about 0.01. To check the quality of image, spatial quality evaluator (BRISQUE) value is measured. From Tables 1 and 2, it is clearly observed that the proposed method provides the better visual enhancement results as the WB-MUV methods failed to provide the perfect enhancement for dusty and low light images. And from Table 1, it is proved that the proposed method provides the better quantitative evaluation compared to the conventional approaches such as conventional with respect to the various metrics such as PSNR, MSE and BRISQUE, respectively.

6 Conclusion

In this paper, the DLCNN algorithm is implemented for enhancing the underwater images by decreasing the depth of the network to 10. A novel architecture is designed for this execution. This design enhances the underwater images more by utilizing contrast limited adaptive histogram equalization algorithm. The multilayer deep learning methods give the extended enhancement, because various layers are used to analyse the each and every pixel region. If the low intensity regions are identified, then they are enhanced by the pre-trained DLCNN network model. The proposed model also responsible for the smoothness enhancement, brightness control, saturation, hue and intensity adjustments to the standard levels. For enhancing the light, edge difference loss levels are calculated. The MEF method responsible for controlling all the parameters, for this purpose, mean preserving brightness levels are developed from the DLCNN method. The feature specific preserving values capable of correcting the all errors with detailed enhancement, by using the deep learning-based brightness preserved histogram specification controlled gives the outstanding results compared to the existing WB-MUV approach. This work can be extended to implement the satellite image processing applications for haze removal.

References

1. Berman D et al (2020) Underwater single image color restoration using haze-lines and a new quantitative dataset. IEEE Trans Pattern Anal Mach Intell
2. Cho Y et al (2020) Underwater image dehazing via unpaired image-to-image translation. Int J Control Autom Syst 18(3):605–614
3. Hou G et al (2020) Underwater image dehazing and denoisingvia curvature variation regularization. Multimedia Tools Appl 79(27):20199–20219 (2020)
4. Liu C, Tao L, Kim YT (2020) VLW-Net: a very light-weight convolutional neural network (CNN) for single image dehazing. In: International conference on advanced concepts for intelligent vision systems. Springer, Cham
5. Pérez J et al (2020) Recovering depth from still images forunderwater Dehazing using deep learning. Sensors 20(16):4580
6. Almero VJD, Concepcion RS, Alejandrino JD, Bandala AA, Española JL, Bedruz RAR et al (2020) Genetic algorithm-based dark channel prior parameters selection for single underwater image dehazing. In: IEEE region 10 conference (TENCON). IEEE, pp 1153–1158
7. Zhu Z et al (2020) A novel fast single image dehazing algorithm based on artificial multiexposure image fusion. IEEE Trans Instrum Measur 70:1–23
8. Kim DG, Kim SM (2020) Single image-based enhancement techniques for underwater optical imaging. J Ocean Eng Technol 34(6):442–453
9. Liu Y et al (2020) Underwater single image dehazing using the color space dimensionality reduction prior. IEEE Access 8:91116–91128
10. Cai C, Zhang Y, Liu T (2019) Underwater image processing system for image enhancement and restoration. In: IEEE 11th international conference on communication software and networks (ICCSN), Chongqing, China, pp 381–387. https://doi.org/10.1109/ICCSN.2019.8905310
11. Khan A, Ali SSA, Malik AS, Anwer A, Meriaudeau F (2016) Underwater image enhancement by wavelet based fusion. In: IEEE international conference on underwater system technology: theory and applications (USYS), Penang, pp 83–88. https://doi.org/10.1109/USYS.2016.789 3927
12. Pizer SM, Johnston RE, Ericksen JP, Yankaskas BC, Muller KE (1990) Contrast-limited adaptive histogram equalization: speed and effectiveness. In: Proceedings of the first conference on visualization in biomedical computing, Atlanta, GA, USA, pp 337–345. https://doi.org/10.1109/VBC.1990.109340
13. Wang Y, Zhang J, Cao Y, Wang Z (2017) A deep CNN method for underwater image enhancement. In: IEEE international conference on image processing (ICIP), Beijing, pp1382–1386. https://doi.org/10.1109/ICIP.2017.8296508
14. Yadav G, Maheshwari S, Agarwal A (2014) Contrast limited adaptive histogram equalization based enhancement for real time video system. In: International conference on advances in computing, communications and informatics (ICACCI), New Delhi, pp 2392–2397. https://doi.org/10.1109/ICACCI.2014.6968381
15. Deng X, Wang H, Liu X, Gu Q (2017) State of the art of the underwater image processing methods. In: IEEE international conference on signal processing, communications and computing (ICSPCC), Xiamen, pp 1–6.https://doi.org/10.1109/ICSPCC.2017.8242429
16. Fu X, Fan Z, Ling M, Huang Y, Ding X (2017) Two-step approach for single underwater image enhancement. In: International symposium on intelligent signal processing and communication systems (ISPACS), Xiamen, pp 789–794.https://doi.org/10.1109/ISPACS.2017.8266583
17. Liu P et al (2019) Underwater image enhancement with a deep residual framework. IEEE Access 7:94614–94629
18. Tao Y, Dong L, Xu W (2020) A novel two-step strategy based on white-balancing and fusion for underwater image enhancement. IEEE Access 8:217651–217670

Machine Learning Approach-Based Plant Disease Detection and Pest Detection System

V. Radhika, R. Ramya, and R. Abhishek

Abstract Agriculture yields account for over 70% of India's GDP. But agriculture and farmers go through a major loss due to plant diseases. The diseases on the plants can be noticed through change in color on the leaves or on the stem of the plants. The diseases in plant occur through fungus, bacteria, and viruses. Usually, the plants are monitored manually at every stage of plant life cycle. But this is time consuming and not much effective. Plant diseases can be easily detected through the advancement in image processing techniques that can provide good results with less human efforts. This will also help to control the plant disease at the beginning stage of infection. Techniques like image acquisition, preprocessing, spot segmentation, feature extraction, and classification were carried for detecting the disease. K-means clustering method and thresholding algorithms were used for image segmentation. The data for various plants with and without diseases were used to train the classifier model. The extracted features from clustering methods are given as classifier inputs to classify the disease. Investigation were carried out on Early scorch, Ashen mold, Late scorch, Cottony mold, and Ting whiteness diseases of plants using K-means clustering.

Keywords Plant disease detection · Classifier · Image processing · K-means · Thresholding

V. Radhika (✉) · R. Ramya · R. Abhishek
Sri Ramakrishna Engineering College, Coimbatore, India
e-mail: radhika.senthil@srec.ac.in

R. Ramya
e-mail: ramya.r@srec.ac.in

R. Abhishek
e-mail: abishek.2006002@srec.ac.in

1 Introduction

Pests and diseases affect the growth of plants that creates a major threat to food production. Plant diseases not only threaten food security, they also affect millions of smallholders. Various approaches are developed in the literature to prevent the crop loss due to the plant diseases [1]. The old and classic approach of plant disease detection is relied on observation with the naked eye, which is time-consuming and inaccurate. In some countries, professional consultants are involved, which requires an investment of time and money. Improper monitoring of various diseases affects plant growth and requires the use of more chemicals to treat disease. These chemicals are toxic to humans, pets, insects, and birds and are beneficial to agriculture. The latest imaging technologies help to solve the problem. Automatic identification of plant illnesses is critical for detecting disease symptoms as soon as they occur on the plant's growing leaf and fruit [2]. This article introduces a MATLAB-based system that focuses on the diseased area of leaves and fruits and uses an image processing technique to accurately detect and identify plant diseases. MATLAB image processing begins by capturing high-resolution digital images saved for experiments [3]. The images are then used for preprocessing to enhance the image. The captured images of leaves and fruits are segmented using the K-means clustering method to form groups. The features are extracted before K-means and the SVM algorithm is used for training and classification [4] and diseases are detected. Insects and pests also create a major loss to human beings, and controlling and maintain their population is very important to ensure the crop health [5]. Insect and pests are considered to be responsible for approximately 10–20% of yield losses occur in crops worldwide. The detection of these insects and pests in food supply chain is the major challenge. Extreme use of pesticides on crops results in severe water and soil contamination. Chemical components retained in the crop also affect the consumer. According to World Health Organization (WHO), every year pesticides affect the life cycle of millions of people. Farmers do not have any knowledge on the usage of pesticides and effect human beings and other organisms [6]. Different methods implemented like glue boards, toxic baits, pesticides, and others were used to kill the pests remotely and effectively, thus increasing productivity and thereby enhancing the economy. The prototype is developed using Arduino technology, which includes specific sensors and a wireless module to collect instant data online.

Integrated pest management is being developed to improve the management of pests, reduce the usage of pesticides, and improve the yield in agriculture. Also the expose of pesticides to the human leads to health issues like reduction in speed of response to stimuli and other effects like allergies and asthma. Recent technology using radar technology for monitoring the pest, video equipment for observing the pests, and thermal imaging techniques and chemiluminescent method for tracking insect movement are available in the literature. These are costly and not affordable for farmers. Existing system is simulation based. With recent advancement in sensor technology, microprocessors, and telecommunication allow for the development of smart system that can be employed in the field that detect and monitor the pest

infestation. This new developed technique can detect and monitor the pest infestation at early stages to avoid worse damage.

Usually, the pest infestation is only recognized in the last stage and the recovery time is almost over, which affects the plants very severely. In this proposed work, pests infestation is detected with sound sensors. These sensors are placed across the farm and pick up sounds in the vicinity. Sound waves from field were analyzed for pest infestation as different pests exist at different range of sound frequency. If the noise level rises above the threshold, the farmer's attention is drawn to the specific area in which the pest infestation occurs. An audio probe is used that has an acoustic sensor for the acquisition of noise. The noise signal can be amplified with an amplifier so that it can be further processed with low power processor. Algorithms that run on processor and a wireless interface can detect the pest remotely. The Internet of Things has brought the huge opportunities for development of science and social economic. This Internet of Things plays an important role for the development of modern agriculture and for the betterment of current agricultural scenario also.

Several sensors placed in the field transmit the noise level to the acoustic sensor nodes in the base station. The base then transmits the information to the computer, which modifies the farmers to take the necessary steps. A central server can also be accessed via web applications in order to call up the details of the pest infestation. This detection helps farmers to control the infestation at a very early stage and thus to reduce the high proportion of crop losses. This monitoring system offers good scalability with low power consumption, which can be implemented in greenhouses and large plantations.

2 Existing Method

Researchers who wished to quantify impacts based on color information employed multispectral or hyper spectral imaging point source devices (such as spectrora-diometers that do not provide a spatial image) to collect data before hyperspectral imaging devices were available. In this work, a pointer is selected and then measuring button is clicked. Instead, it is the user's obligation to create the capture procedure. The massive numerical datasets that result must then be evaluated to offer valuable information. It observes changes in circumstances at selected key spots in the spectrum from a small number of sites in the wavelength range. This method is also used to mitigate the impacts of relative change. This includes a combination of two or more wavelengths, commonly known as a "list". (A list containing, can be further determined in multispectral by the satellite). In connection with plant tissue, these records are called "vegetation lists" and are either general characteristics of the plant or specific parameters of its development [7].

One of the most popular and widely used metrics is the Normalized Differential Vegetation Index (NDVI), which is used to measure the general health of a crop calculated with a simple relationship close to IR and visible light. The NDVI has been used for many different purposes, such as detecting stress from the Sunn pest/grain

pest Eurygaster integriceps put. (Hemiptera: Scutelleridae), in wheat. Most indexes are very specific and only work on the datasets for which they were designed [8]. There are disease-focused studies that focus on developing disease indices to identify and quantify specific diseases, for example one study used the Rust Severity Index on leaf rust (LRDSI) with an accuracy of 87–91% to detect leaf rust (Puccinia triticina).

3   Plant Leaf Disease Detection System

The proposed step-by-step approach is to collect the leaf image database, preprocess these images, segment these images using the K-means clustering method, extract features using the GLCM method, and finally, train the system using the SVM algorithm. First step in the process is to train the system with the healthy and unhealthy leaves. Images of different types of leaves are obtained from the Internet. Various steps of image processing techniques are used to process these images. By processing these images, various and useful features for analyzing the leaves are obtained.

Algorithm shown below illustrates the step-by-step approach for the proposed image recognition and segmentation process.

 (i)   Images acquisition is the first step to get the healthy and unhealthy leaves.
 (ii) The input image is preprocessed in order to improve image quality and remove unwanted image distortion. This step preserves the relevant area of the image and smoothes it. Image and contrast enhancement are also performed.
 (iii) Following that, image segmentation is performed to extract the image's shape, texture, and color features.
 (iv) The next step is feature selection, which is done based on the correlation of the variable and the target variable.
 (v)   Following feature selection, the data is fed into the classifier, which searches for patterns in the data. Classifiers are used to identify various diseases that appear on plant leaves.

The process done for plant disease classification in each step is explained below.

3.1   Image Acquisition

The database collection is the basic step to follow. It is the process of acquiring the images through the camera or from image databases. Raw images related to plant diseases are collected. The image database is very important for the better classification in the detection phase.
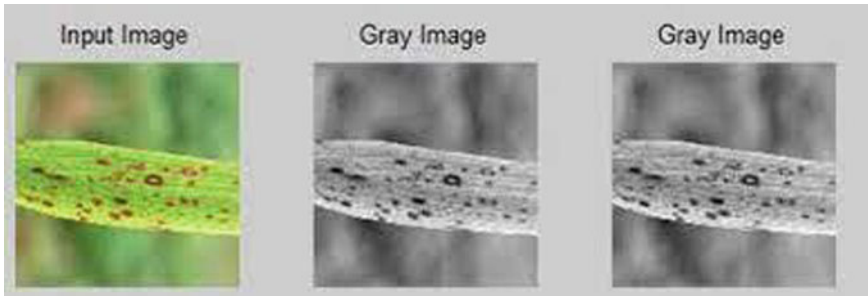
Fig. 1 Contrast enhanced and RGB to gray converted image

3.2 Image Preprocessing

Image preprocessing is done to get the precise results; some background noise is removed before the extraction of features. The RGB image is converted to grayscale, and the Gaussian filter is used for smoothening the image. In this work, images are resized to improve the contrast and the conversion of RGB to grayscale is also done as shown in Fig. 1 for additional operations such as creating clusters in the segmentation.

3.3 Image Histogram

The histogram is created by looking at all of the objects in the image and designating each one to a bin based on its intensity. The image size given to a bin determines its ultimate value image histogram is shown in Fig. 2.

3.4 Image Segmentation

Image segmentation is done to remove the unwanted objects or background that surrounds the image. It partitions the digital in to various parts to derive the focused part from the image. In segmentation, K's clustering method is to subdivide the images into groups where at least part of the group contains an image with the main region of the diseased organ. K-means clustering algorithm is used to classify objects into K classes based on the set of features. The classification is done by minimizing the sum of squares of the distances between the data objects and the corresponding group. The image is converted from RGB color space to $L * a * b *$ color space, where $L * a * b *$ space consists of a luminance layer $L*$, a color saturation layer $a*$ and $b*$. All color information is at the $*$ and $b*$ levels, and the colors are arranged
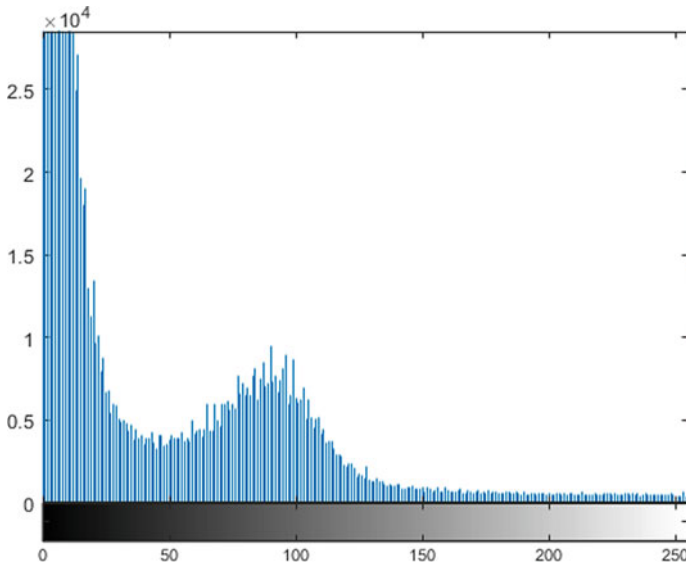
Fig. 2 Representation of image histogram

in K-means groups in the $* b *$ space. Each pixel in the image is tagged from the K-means results and segmented images with disease are generated. In this experiment, we will use segmentation technique to divide the input image into three groups to get good segmentation result. Segmentation of leaf image with three groups formed by K-means clustering method is shown in Fig. 3.
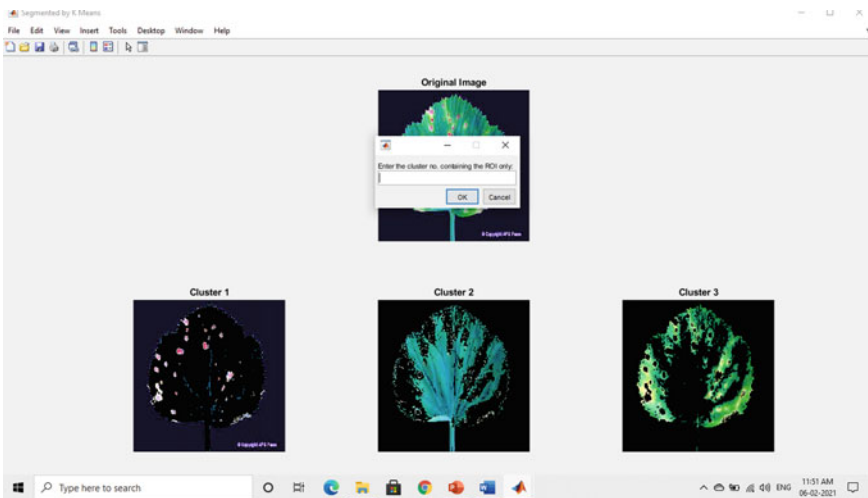
Fig. 3 Diseased leaf image clusters

3.5 Feature Selection

After applying the image preprocessing and image segmentation on the disease infected leaf, feature extraction techniques are applied to obtain the targeted features for disease classification. The overall accuracy of machine learning algorithm depends on this feature selection step. The features are generally classified into categories like (i) color feature, (ii) texture feature, (iii) shape feature, (iv) vein feature, and (v) moment feature.

The statistical properties of the texture are obtained using the Grayscale Interaction Matrix (GLCM) formula for texture analysis, and the texture properties are calculated from the statistical distribution of the observed intensity combinations at a given location. The order of the first, second, and highest number of intensity points in each combination. Various statistical properties of GLCM textures are energy, sum of entropy, covariance, measure of correlated entropy information, contrast, inverse difference, and difference.

3.6 Classifier

In classification step, the data is separated in to training dataset and test datasets. The training set consists of target values and attribute values for each data. A hyperplane is located that split the points into two different classes as positive plane and negative plane. The type of disease is identified through this step and intimated to farmers.

4 Pest Detection System

An acoustic sensor is used to monitor the noise level from pests and every time the noise exceeds the threshold, it alerts the farmer to the area of the pest infestation, introducing automation in the field of agriculture and reducing the farmer's efforts. The noise level throughout the operation is recorded by a wireless sensor node that is placed above the field. The noise level in digital format is sent to the processor [9]. If the noise level exceeds the threshold value, the information is transmitted to the farmer via a radio module on his mobile phone [10]. The block diagram of the pest detection system is shown in Fig. 4.

5 Results and Discussion

The classification is initially based on the K-mean clustering minimum distance criterion and shows efficiency with an accuracy of 86.54%. The detection accuracy
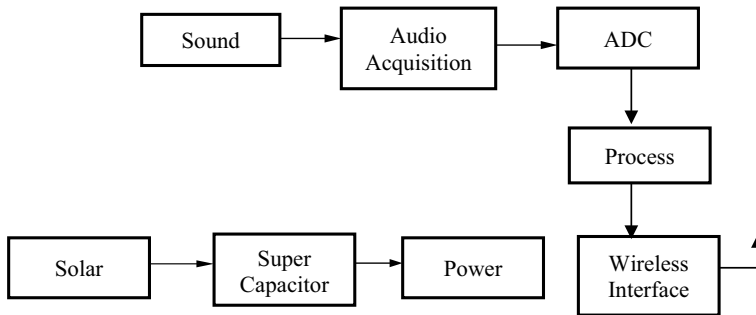
Fig. 4 Pest detection system

was improved to 93.63% by the proposed algorithm. In the second step, classification is performed using an SVM classifier, which shows efficiency with an accuracy of 95.71%. Using the proposed algorithm, the SVM detection accuracy was improved to 95.71%. The detection accuracy using SVM can be seen from the improved results using the proposed algorithm compared to other claimed approaches. The results of the plant disease detection system are shown in Fig. 5.

6 Conclusion

In this work, a system for the detection and monitoring of insect numbers was successfully implemented, which was able to test and inform possible causes of the infestation of insect pests in cultivated plants and to know the concentration of insects in a greenhouse. The results of this work can be used as a reference for future applications, such as the early prevention and prediction of insect pests, not only for green house applications, but also for field applications from the system.

This disease identification method provides an efficient and accurate method for detecting and classifying plant diseases using MATLAB imaging. The methodology proposed in this article depends on the K-means and MultiSVM methods configured for both server blades. MATLAB software is ideal for digital image processing. K-means clustering algorithm and SVM are very accurate and take very little time to complete the processing.
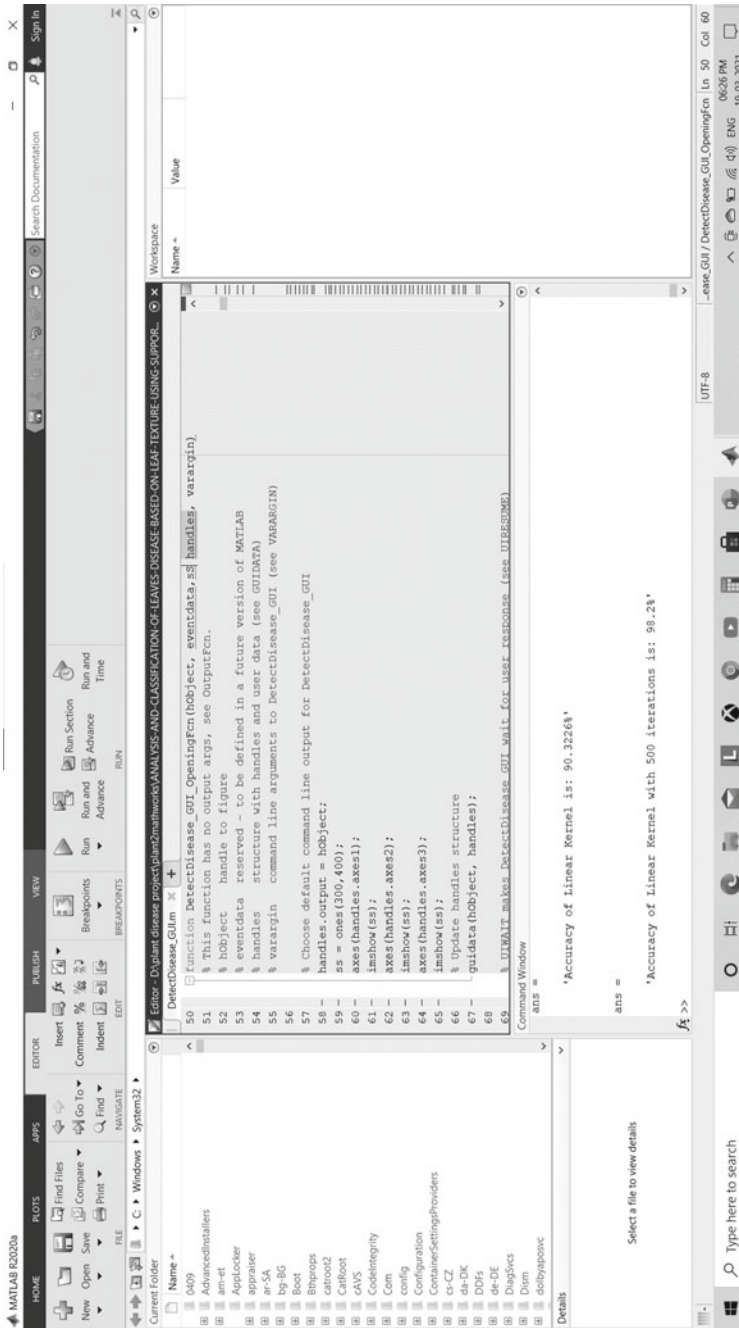
Fig. 5 Result of plant disease detection

References

1. Elijah O, Rahman T, Orikumhi I, Leow CY, Hindia MN (2018) An overview of Internet of Things (IoT) and data analytics in agriculture: benefits and challenges. IEEE J Internet Things 5:3758–3773
2. Pajares G (2011) Advances in sensors applied to agriculture and forestry. J Sensors 11:8930–8932
3. Mattihalli C, Gedefaye E, Endalamaw F, Necho A (2018) Plant leaf diseases detection and automedicine. J Internet Things 1:67–73
4. Yogamangalam R, Karthikeyan B (2013) Segmentation techniques comparison in image processing. Int J Eng Technol 5:307–313
5. Vinayak SV, Apte SD (2017) Wireless sensor networks and monitoring of environmental parameters in precision agriculture. Int J Adv Res Comput Sci Software Eng 7:432–437
6. Gasso-Tortajada V, Ward AJ, Mansur H, Brøchner T, Sørensen CG, Green O (2010) A novel acoustic sensor approach to classify seeds based on sound absorption spectra. Sensors 10:10027–10039
7. Sun Y, Liu Y, Wang G, Zhang H (2017) Deep learning for plant identification in natural environment. Comput Intell Neurosci. https://doi.org/10.1155/2017/7361042
8. Joly A, Goeau H, Bonnet P (2014) Interactive plant identification based on social image data. Ecol Inf 23:22–34
9. Venkatesan R, Kathrine G, Jaspher W, Ramalakshmi K (2018) Internet of Things based pest management using natural pesticides for small scale organic gardens. J Comput Theor Nanosci 15:2742–2747
10. Shi J, Yuan X, Cai Y, Wang G (2017) GPS real-time precise point positioning for aerial triangulation. J GPS Solut 21:405–414

RETRACTED CHAPTER: Design of Low Voltage Pre-settable Adiabatic Flip-Flops Using Novel Resettable Adiabatic Buffers

Divya Gampala, Y. Prasad, and T. Satyanarayana

The editor has retracted this article because of significant overlap with a previously-published article by different authors [1]. All authors agree to the retraction. [1] S. Maheshwari, V. A. Bartlett and I. Kale, "Adiabatic flip-flops and sequential circuit design using novel resettable adiabatic buffers," 2017 European Conference on Circuit Theory and Design (ECCTD), Catania, Italy, 2017, pp. 1–4, doi: 10.1109/ECCTD.2017.8093257

D. Gampala (✉) · Y. Prasad · T. Satyanarayana
Department of ECE, CMR Engineering College, Medchal, India
e-mail: dha.grs@gmail.com

Y. Prasad
e-mail: yarrabadiprasad@cmrec.ac.in

T. Satyanarayana
e-mail: satyant234@gmail.com

Raspberry Pi-Based Smart Mirror

K. Subramanya Chari, Mandapati Raja, and Somala Rama Kishore

Abstract With the recent launches of innovative and smart goods that act like computers and mobile phones, the market is booming. Salon mirrors, as well as 3D mirrors, are substantially more expensive, with public use restrictions. This project uses the Raspberry Pi to create a home-based Internet of Things (IoT). A smart mirror is proposed, with a Raspberry Pi as the host controller and an STM32F030C8T6 microcontroller as the control chip at the core level. The system attached to the Raspberry Pi is connected to the Wi-Fi network via API, which shows weather data, time, and date on the display. The user can ask the APP mirror for any information, such as the time, weather, and other facts, and the mirror will broadcast the requested information. The proposed smart mirror is a low-cost gadget with a low level of complexity and a compact size that may be used for a variety of purposes.

1 Introduction

The standards and the quality of the life are improving day by day because of upcoming embedded devices using various technologies. By interactive computing, we can achieve various benefits by integrating various technologies. We can design various innovative products or devices by means of wireless networks.

It provides not only comforts to the one who uses it, but also security and convenience to various users at homes and industries in various ways. Currently, various devices like smart watches and smart TVs already existed and now comes the new concept which is smart mirror. These smart mirrors are nothing but intelligent mirrors that are limited to few places like public places and commonly used in hair salon shops in foreign countries.

K. Subramanya Chari (✉) · M. Raja · S. R. Kishore
CMR Engineering College, Hyderabad, India
e-mail: chari451@gmail.com

The reason why it is not used at public places is that they are expensive, and there is slight delay in the picture. Daily we spend some time after we fresh up to find out how our dress is looking, get ready by looking at it. This time can be effectively used for getting latest updates regarding weather, news, time, and other information too. Lot of time can be saved by using these intelligent mirrors. All the above-listed smart features like providing weather report, local time, and other data can be obtained by using Raspberry Pi.

For dealing with such situations, we can control the electrical appliances at home by providing convenience to all the users of it by means of network connection between the devices and the lamps respectively. It works in the following way like firstly the users need to give certain instructions to it and the sensors present in the system will be able to recognize their voice and respond accordingly based on the requirement.

For this purpose, we design a system that is based on three main objectives. In which the first objective is designing an intelligent mirror that is a smart mirror based on Raspberry Pi, secondly, for implementing the smart mirror, we need a voice recognition system, and finally, we need to perform the testing on the Raspberry Pi for checking how it works for users.

2   Related Work

The smart mirror which is presented in this paper provides services to the users by facilitating access to them in various ways as mentioned. It basically acts as an interface between the devices naturally. This interface is used in the environment for interaction as commented below.

The well-known context-aware home environments are Philips Home Lab [1]. They provide various projects to the users as per their preferences. Their works are of intellectual and creative. Let us take for instance the interactive mirror [2] by them. They are presented in the bathrooms and provide some information according to the user's interests. Here are the few activities that can be performed by the children. They can spend time in watching their favorite videos while they are brushing. Even elders can spend time for knowing weather reports of various places, watch interesting ones. Here both the LCD's flat screens are combined together by the help of central processor for facilitating various intended services to the users according to their interests.

Smart environments are realized by the use of ambient artificial intelligence (AMI) [3]. It results in the vast changes in the domain like industry and home environments by its intelligence The change brought by the AMI is that humans are surrounded by various interfaces which are connected together, and it is given by European consortium. It is very user friendly for interacting with humans.

When this AMI is employed in the home environment, it provides quality, convenience, as well as the security to the residents by providing utmost safety for both elderly people and people with disabilities and physically challenged people at home, hospitals, and workplaces. It is mostly helpful for people with disabilities.

The innovations of AMI have created a surprise to various fields like home automation and workplaces. The regular mirror which we keep at homes does not do any smart work. We waste around twenty minutes or so in front of the natural mirror. So to overcome the wastage of time, we designed a prototype called smart mirror that covers the time we spend some time in reading newspaper. It can be viewed as the problem statement for the paper.

3 Proposed Work

In proposed work, we designed smart, intelligent mirror which uses Raspberry Pi and STM32 controller. Smart mirror is the combination of Raspberry Pi, controller module, Wi-Fi module, and API interface. Controller module is the core brain of the proposed mirror which is the main part of the system. It consists of two modules in proposed work: one is microcontroller STM32 and another is Raspberry Pi module.

Display module consists of plasma display which is used to see the functional changes with size of 24 inch and one-way display. Clock module is used for getting exact time and date with the help of CMOS-based module. For wireless transmission and reception, we used Wi-Fi module. For hardware and software module connectivity, we used programming.

As shown in Fig. 1, the proposed work block diagram consists of following modules:

1. Raspberry Pi 3
2. Wi-Fi module
3. IoT module
4. Micro-SD card
5. Power supply
6. Plasma panel display.

In this project, we used Raspberry Pi (Fig. 2) connected to computer and remaining application is implemented. Raspberry Pi is the module which works as small computer which is connected to computer to understand processing using mouse and keyboard.
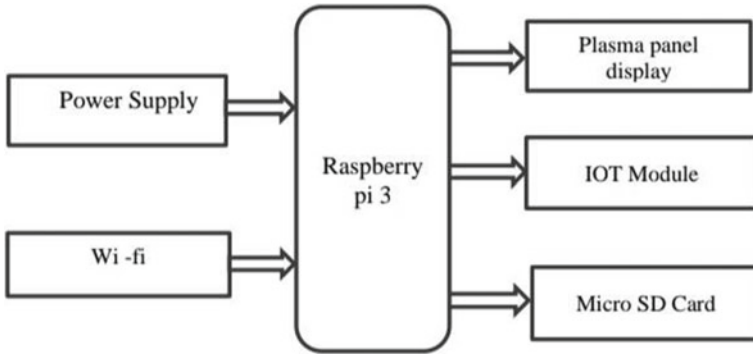
Fig. 1 Proposed work block diagram

Fig. 2 Raspberry Pi module
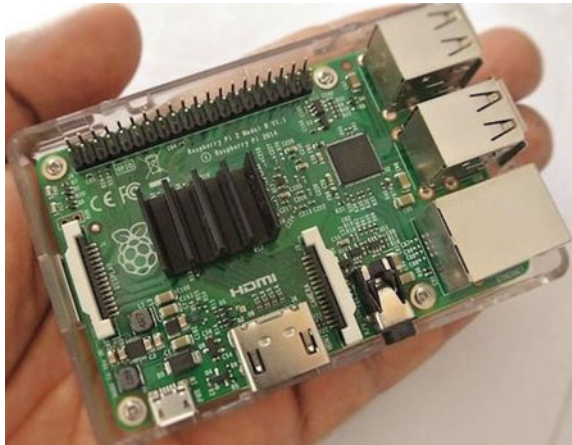
We used ESP Wi-Fi module (Fig. 3) which is a very low-cost module produced by Shanghai-based company for full TCI/IP stack. This module is mainly used to design IoT applications for embedded system. To communicate with this Wi-Fi module, microcontroller sets some commands, and with UART, we need to set some baud rate for communication.

Internet of Things (IoT) is used for better life of human beings with added intelligence. IoT module is designed for wireless communication with flexibility and long duration access. It makes wireless connections between devices connected to system. Depending on the type of application, IoT module will be selected which provides specific features. Micro-SD card is abbreviated as secure device which is designed by SDA (SD Association). It is mainly used in portable devices. Some companies like SanDisk improved multimedia cards (MMCs) to get industry standards. It is used to store information or commands.

Fig. 3 Wi-Fi module
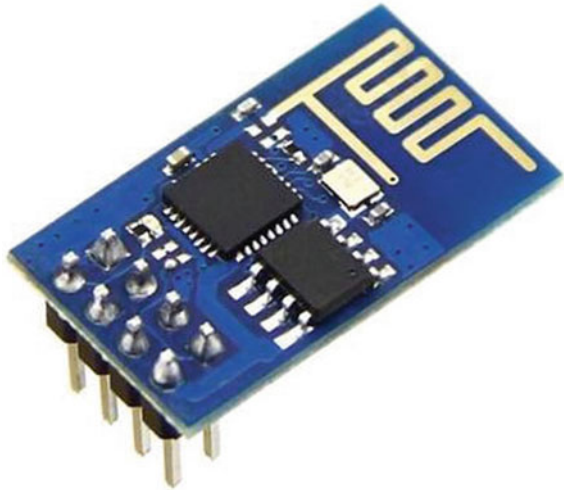
Power supply is the main requirement for working of this proposed device. There are multiple parts for proposed module which require different voltage and current requirement. Raspberry Pi module requires 5 V input, while other devices may work with 230 V input.

We used plasma panel display, which uses plasma, ionized gas that respond to electrical field. It is very flat panel display. Previously, it is used in television (TV). This can be used for smaller application as for big application in the market LCD and LED took very much attention of users than plasma display.

The working flow starts with initialization of Raspberry Pi model. In second step, the process includes two steps: One is collecting the data from STM32 [4] like temperature, time, date, etc., and second is getting information from API like weather, etc. (Fig. 4).

Many other information is required while designing like detection of anyone in front of door or no, is anyone press any button, depending on that there is need of response from smart mirror as output.

Controller STM32 is used for controlling commands for interface. STM32 is a 32-bit microcontroller and designed by STMicroelectronics. There is need of initialization, reading an information such as time, date, and temperature and sending information to Raspberry Pi through serial port. From Raspberry Pi, further process will be considered [5] (Fig. 5).

4 Results

Below, complete hardware output information is mentioned which we got from smart mirror device implemented with the help of Raspberry Pi [6].
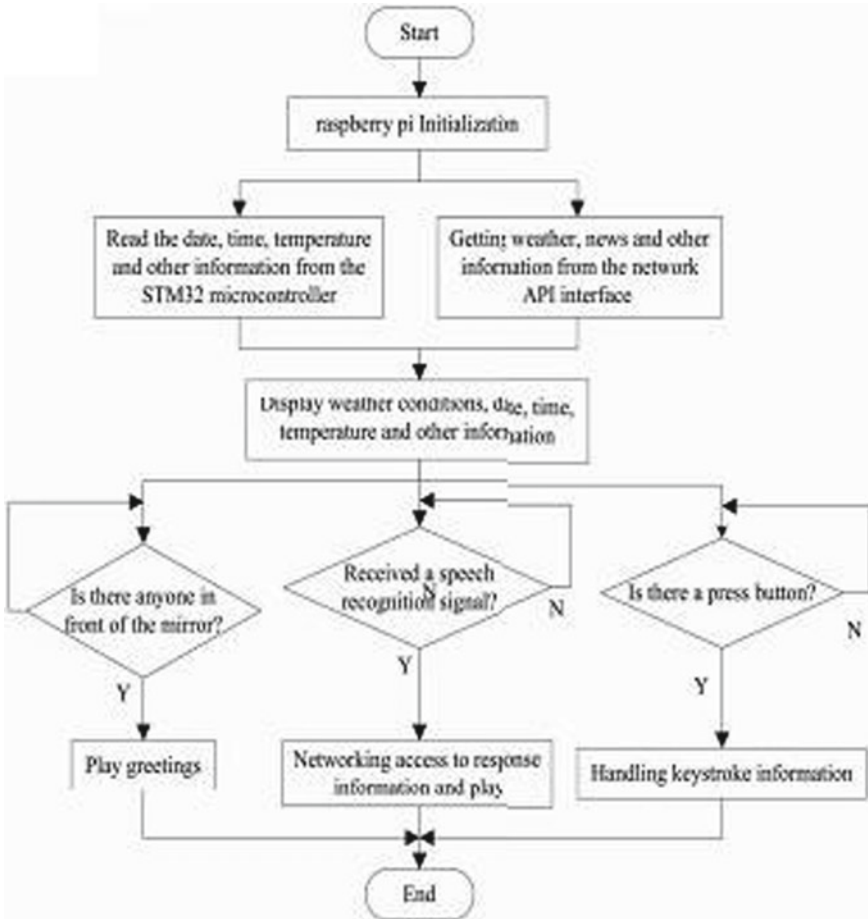
Fig. 4 Raspberry Pi working flowchart

Proposed hardware connection between Raspberry Pi, microcontroller, and other devices is shown in Fig. 6. Hardware connections are very important to get proper transfer of information from one module to other and to get perfect output at last [7].

Smart mirror displays the project title, temperature, and time on plasma display (Fig. 7). This display is made of plasma display which is preferred over LCD and LED for our application.

At morning time, smart mirror displays 'Good Morning!!' message and news information at the bottom (Fig. 8).

Smart mirror is showing weather information on plasma display. It is showing the specific day as well as 1 week weather information on display (Fig. 9).

Fig. 5 Controller STM32 working flow chart
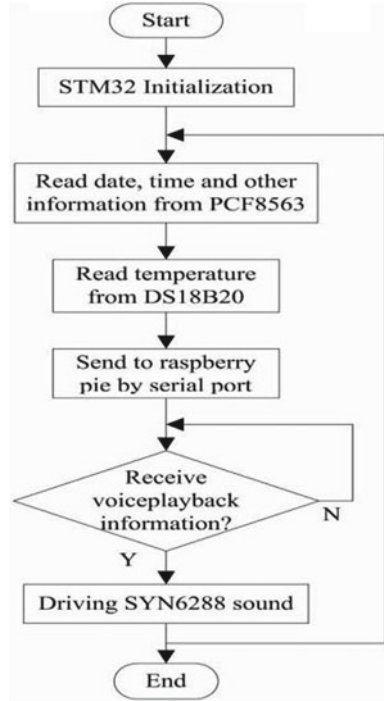
Fig. 6 Proposed smart mirror interface with mobile
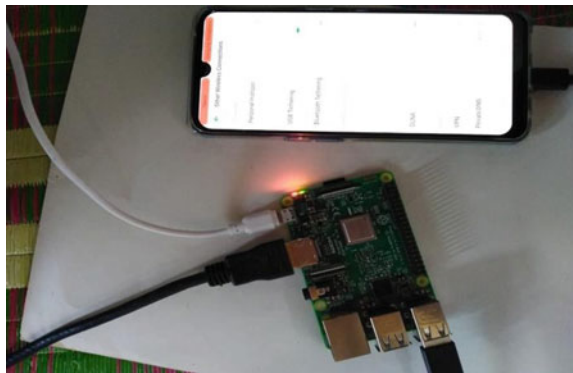
5 Conclusion

In this application, we used host controller as Raspberry Pi and main controller as microcontroller STM32. Raspberry Pi and Wi-Fi module are interlinked which provides weather details, date, time, etc., with the help of API interface which is further showed on plasma display for users. User can very easily interact with proposed smart mirror using API interface from the mobile app. Smart mirror is

Fig. 7 Main display for smart mirror system
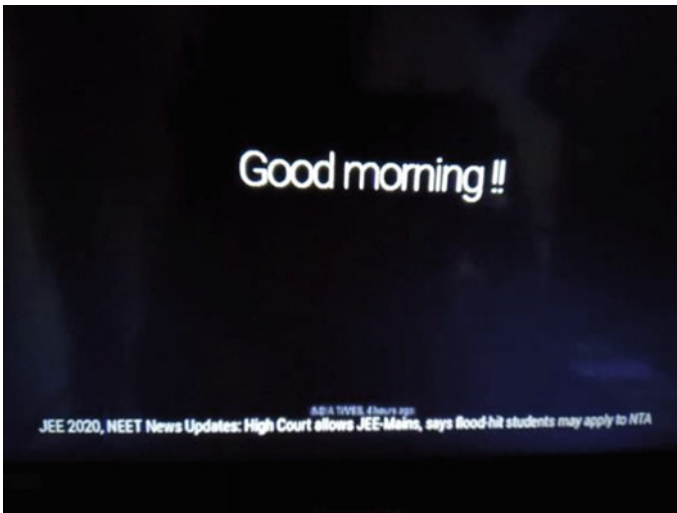
Fig. 8 'Good Morning!!' display on smart mirror

preferred over other applications as it has very simple operations with small and low-cost hardware part. It is so much user friendly for even new user. Smart mirror has many advantages in different domain as we as family purpose can use it. Further same model can be used for face recognition, object identification, etc. In the year 2020, all the special days will be displayed. Using this display, the nearest special day can be seen on display.
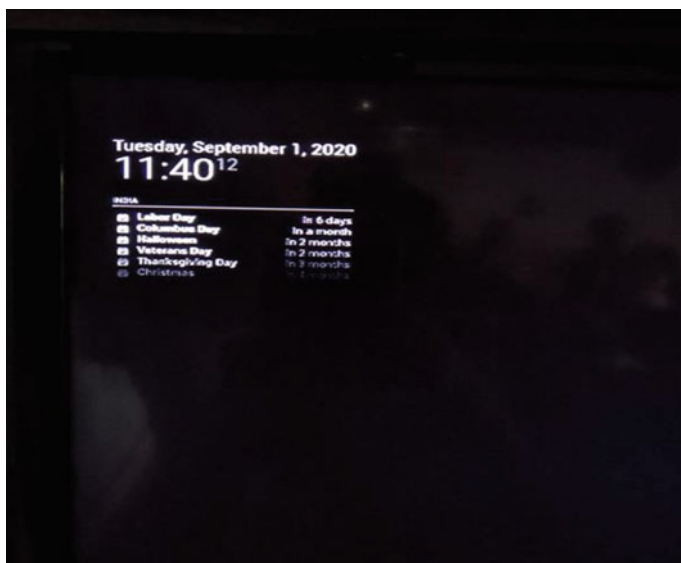
Fig. 9 Smart mirror display showing special days
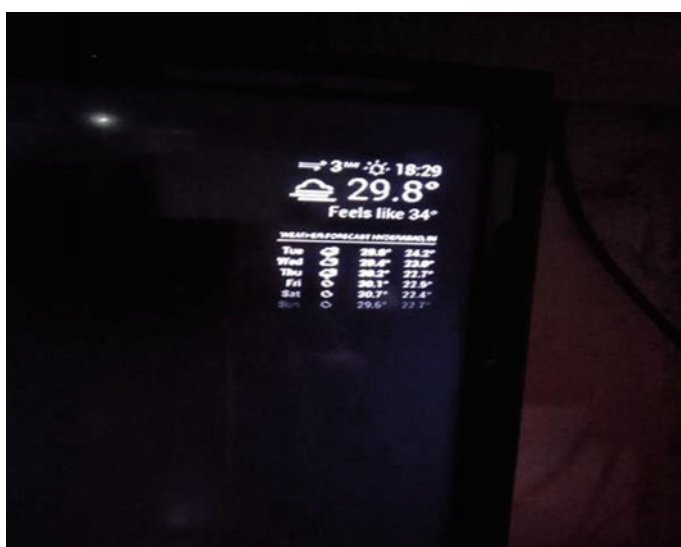
Fig. 10 Smart mirror display showing weather details

References

1. Chaitanya U, Sunayana KV (2020) Voice assistant and security-based smart mirror. Int J Recent Technol Eng (IJRTE) 8(6)
2. Pant AP, Naik DS, Dandgawhal TP, Patil SR, Kapadnis JY (2017) IOT based smart mirror using credit card sized single board computer. IJARIIE 3(2)
3. Katole M, Khorgade M (2018) A novel approach to designing a smart mirror with raspberry pi. Int J Eng Technol Sci Res 5(3)
4. Mittal DK, Verma V, Rastogi R (2017) A comparative study and new model for smart mirror. Int J Sci Res Comput Sci Eng 5(6):58–61
5. Maheshwari P, Kaur MJ, Anand S (2017) Smart mirror: a reflective interface to maximize productivity. Int J Comput Appl 166(9)
6. Akshaya R, Niroshma Raj N, Gowri S (2018) Smart mirror digital magazine for university implemented using raspberry pi. In: International conference on emerging trends and innovations in engineering and technological research (ICETIETR)
7. https://howchoo.com/g/ntcymzbimjv/how-to-installmagic-mirror-on-your-raspberry-pi/how-to-installmagic-mirror-on-your-raspberry-pi

Deep Learning Framework for Object Detection from Drone Videos

Somala Rama Kishore, K. Vani, and Suman Mishra

Abstract The extensive research in the space of deep learning has resulted in unimaginable solutions to various problems around us. Deep learning has spread its roots fairly into computer vision, speech, and text. Vision applications through CCTVs, mobiles, and cameras are dominating the market. Hence, this paper attempts to display that aerial-based videos which are possible to fetch through drones when coupled with computer vision-based object detection deep learning networks will give promising results leading to remarkable solutions benefiting the sectors of agriculture, surveillance, military, logistics, search, and rescue. Furthermore, drones have less battery and computational hardware; therefore, performing object detection on drones is highly challenging. It is very essential to understand how an object detection model would perform on certain architecture before deploying the model on to drone. Therefore, this paper implements object detection models on a drone-based images and makes an analysis that helps in choosing a best model to deploy by making a performance comparison.

Keywords Deep learning · TensorFlow · Python · Object detection · Drones · Performance analysis

1 Introduction

World has witnessed an exponential increase in the generation of data. Data takes structured and unstructured forms and is being produced continuously from various sources such as sensors, cameras, mobiles, computers, and web. This modern digital era enables machine learning and deep learning to take advantage and leverage this data with neural network architectures to create a significant impact in almost all practical applications. Deep learning is an implementation branch of Artificial Intelligence, where the power lies with computers to mimic the human behavior in the form neural networks by learning insights from data and taking results correspondingly.

S. R. Kishore (✉) · K. Vani · S. Mishra
Department of ECE, CMR Engineering College, Hyderabad, India
e-mail: ramki430@gmail.com

It is evident that deep learning has already created and taken its position in almost every application. According to Hordri et al. [1], few of the applications where deep learning is extensively applied are computer vision, natural language processing, recommendation systems, medical domain for drug discovery, and bioinformatics. However, this list is not exhaustive. Computer vision is one domain where industry is adapting and implementing applications around it at unprecedented pace. Wu et al. [2] explains the application of computer vision-based deep learning in autonomous driving, robotics, retail video analytics, and image analytics. Object Classification, Object Detection, Object Segmentation, and Tracking [3] are few variants of implementing deep learning models around computer vision applications. From these various applications, it is understood that the power of deep learning coupled with imagery or video data is unimaginable. Hence, this paper proceeds to implement object detection models on aerial-based drone views to create remarkable solutions that benefit the sectors of agriculture, surveillance, highways, traffic, etc. Further, this paper develops a method for performance analysis of object detection inference that allows user to choose which model to be used on drone for deployment. This paper is organized into different sections, namely Sect. 2 briefs on our literature survey, Sect. 3 explains the adopted methodology, Sect. 4 demonstrates object detection results, and Sect. 5 discuss the analysis of performance.

2   Literature Survey

Object detection system uses the trained knowledge in identifying and locating objects in images and videos. The localization of objects in images is represented by rectangles which are called as "bounding boxes." More specifically speaking, each object is surrounded by a bounding box depending on the deep learning model and its dataset. Object detection models are powerful than classification models because their localization capability resulting in allowing a user to develop variety of uses cases. This section summarizes significant previous research that had been focusing on computer vision-based object detection. In majority of previous findings, the works were revolving around proposing the architecture of various object detection models and their usage on detection of generalized objects in normal images which are generated from CCTVs, hand-held cameras, mobiles, etc. [4] proposes YOLO and [5] introduces Faster R-CNN for real-time object detection of vehicles such as cars and motors, trees, pedestrians, etc. Efficient CNNs for embedded and mobile-based vision application in the name of Mobile-Nets were introduced in [6]. Besides these, there is much other architecture for object detection depending upon the application which is to be built. However, there were relatively few findings with respect to object detection on drone-based views compared to object detection on a generic view as detecting objects in drone images is challenging because of the orientation and position of objects.

3 Methodology

The paper adopted two phases for our study and experiments of deep learning models. They are as follows

(i)   Training phase. (Or Initial)
(ii) Inference phase. (Or Final)

Each phase is summarized in the following subsections.

3.1 Training Phase

Every deep learning application includes modeling and adjusting the weights across the network based on the use case. In our study, the training phase, which is also called as initial phase summarized in Fig. 1 and followed by explanation of each sub-phase.

1. *Data Accumulation*
   Data is the primary parameter in deep learning applications and is given highest priority. For our study, data takes image form captured from drones. Standard datasets of drone images are minimal. Hence, the paper had accumulated images from multiple sources such as VisDrone [7], Kaggle [8], and YouTube for training and inference purpose. A total of 6471 images were used for training the object detection models for our implementation.

Fig. 1 Block diagram of training phase
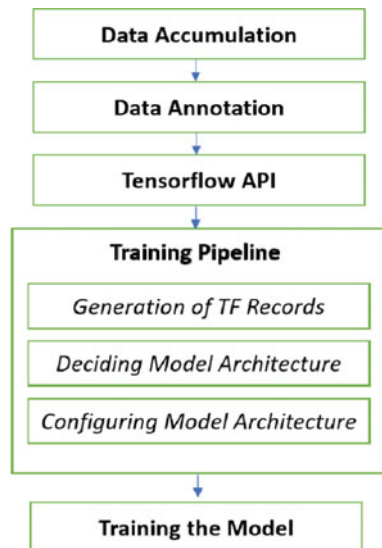
2. *Data Annotation*

   Labeling the data by drawing bounding boxes around each notable object is called as data annotation. Top left *X*, *Y* and bottom right *X*, *Y* are the coordinates which will be noted. Our work had used a graphical image data annotation tool named Labeling [9]. Post annotating, each image has annotations in their corresponding XML file. In our experiment, we have used seven labels, namely pedestrian, bicycle, car, van, truck, bus, and motor.

3. *TensorFlow API*

   TensorFlow is a deep learning framework which is developed by Google. It has got a wide variety and support for multiple models across classification, detection, and segmentation of models. In addition to this, TensorFlow also provides workflows to develop and train the models. We have decided to use TensorFlow because of its wide popularity, ease of use, and extensive capabilities.

4. *Training Pipeline*

   This phase is a critical block in the entire flow of implementation. Our study had adopted three sub- phases as shown in Fig. 1.

   Generation of TF Records: TensorFlow records are the file formats which store sequence of binary records that are suitable to run for TensorFlow runtime. We have created TF records for the images and XML files to initiate training process.

   Deciding Model Architecture: Deciding the number of neurons, hidden layers, type of neural network operations, and output layer neurons are other important decisions for implementing a model. This is highly dependent again on the use case. The state-of-the-art models with a wide range of architectures are fortunately provided by TensorFlow. TensorFlow has a collection of deep learning models in a hub, which is called TensorFlow Zoo. For our study on drone images-based object detection, we had proceeded and decided to use "*Faster* R-CNN Inception ResNet V2" and "SSD ResNet152 V1 FPN" because of its deep architecture.

   Configuring Model Architecture: *Faster* R-CNN Inception ResNet V2 and SSD ResNet152 V1 FPN has a complex architecture and deep operations. These models are trained on COCO dataset. Hence, these are called pre-trained models. In our experiments, we had fine-tuned the network architecture by applying the method of transfer learning. Fully connected layer is configured with our labels.

5. *Training the Models*

   A total of 6471 images of our dataset with a batch size of 32 for almost 30,000 + steps were configured to train both the models. The pre-trained weights of the final layers are completely readjusted according to our dataset. We had used Tesla T4 GPU for training the models. Training took almost 16 + hours. Post training and model checkpoints are saved to use for the inference of images, in the next phase.
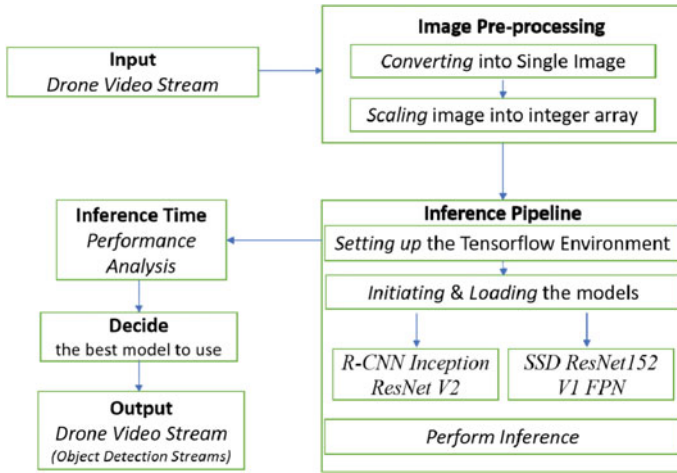
Fig. 2 Block diagram of inference phase

3.2 Inference Phase

Inference phase involves the post-training activities as summarized in Fig. 2. This phase involves preprocessing, inference, and analysis in a high-level view.

1. *Input*
   In our experiments, we had used drone captured images and videos as input to the new re-trained and adjusted models obtained from training phase as discussed above.
2. *Image Pre-processing*
   Videos are sequences or collection of images in a continuous form. Hence, if the input is a video, the paper had processed video into a list of images. If the input is an image, it is directly appended to the list. After this initial step, each collected image is converted into multi-dimensional integer array.
3. *Inference Pipeline*
   This pipeline is critical in inference phase and includes sub-phases. TensorFlow environment was set up. Post setting up the environment, the two models obtained from training phase were initiated and loaded to perform the inference on input.
4. *Inference Time Analysis and Deciding the Model*
   We introduce a method of measuring the inference time taken for one image by computing the difference of start time and exit time of image in and out from the models. We made a comparison and conclusions for deciding the good model based on inference times taken by the two models in the later section. The model which takes less inference time can be used for drone-based deployments excluding the fact of accuracy.

Fig. 3 Input image

5. *Output*

Bounding boxes were sketched on to the input streams. Strictly speaking, on a successful inference the coordinates of each object from our label set is now available for making further use cases.

4 Results

The trained models were set to a confidence score of 0.6 while drawing the bounding boxes on the image from the results obtained from the models' outputs. Few of the sample input and output inferred images from our experiments by the custom trained Faster R-CNN Inception ResNet V2 is in the range of 1.2–1.6 s, and the average time taken by our custom SSD ResNet152 V1 FPN are displayed in Figs. 3, 4, 5, 6, 7 and 8.

5 Conclusion

Our experiments carried out time analysis calculation on individual drone-view images and videos to evaluate the performance of our models. For the images, few drone-viewed images were considered for testing. The average time taken by our custom trained Faster R-CNN Inception ResNet V2 was in the range of 1.2–1.6 s and the average time taken by our custom SSD ResNet152 V1 FPN was in the range of 0.8–1.1 s. The time taken calculation is inclusive of time taken by preprocessing and postprocessing of image. Four fully drone-viewed videos with a total number of 457, 576, 1130, and 1182 frames were considered for testing. The total time taken and average time (single frame of video) taken for inferencing by models are

Fig. 4 Output image by custom Faster R-CNN Inception ResNet

Fig. 5 Output image by custom SSD ResNet152 V1 FPN

Fig. 6 Input image

Fig. 7 Output image by custom Faster R-CNN Inception ResNet V2

Fig. 8 Output image by custom SSD ResNet152 V1 FPN

given in Table 1. Yet again, time taken for calculation is inclusive of time taken for preprocessing and postprocessing of the frames.

Table 1 Analysis of the average time taken for inferencing the frames in the videos

Number of frames	Total time (in sec)		Average time (in sec)	
	Custom FRCNN	Custom SSD	Custom FRCNN	Custom SSD
457	413.46	342.75	0.902	0.75
576	468.99	434.24	0.812	0.752
1130	918.45	733.29	0.812	0.648
1182	962.15	759.53	0.813	0.642

References

1. Hordri NF, Yuhaniz S, Shamsuddin SM (2016) Deep learning and its applications: a review. In: Postgraduate annual research on informatics seminar
2. Wu Q, Liu Y, Li Q, Jin S, Li F (2017) The application of deep learning in computer vision. In: Chinese automation congress (CAC)
3. Yilmaz A, Javed O, Shah M (2006) Object tracking: a survey. ACM Comput Surv 38(4)
4. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection
5. Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6)
6. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) MobileNets: efficient convolutional neural networks for mobile vision applications
7. VisDrone is a dataset which includes images and videos taken by drone. https://github.com/VisDrone/VisDrone-Dataset
8. Kaggle is a dataset resource which includes data in all forms. https://www.kaggle.com/kmader/drone-videos
9. LabelImg is an annotation tool for drawing bounding boxes around the objects. https://github.com/tzutalin/labelImg

RETRACTED CHAPTER: VoteChain: Electronic Voting with Ethereum for Multi-region Democracies

Vaseem Ahmed Qureshi, G. Divya, and T. Satayanarayana

The Editor has retracted this article because its content has been duplicated from an unpublished manuscript authored by Gandikota Ramu, Mohammed Tayeeb Hasan, Krishna Chandra Boyapati and Ali Affan without permission. All authors agree to this retraction.

V. A. Qureshi (✉) · G. Divya · T. Satayanarayana
ECE Department, CMR Engineering College, Medchal, Hyderabad, India
e-mail: qureshi.vaseem@gmail.com

Design and Analysis of Multiband Self-complimentary Log Periodic Tripole Antenna (LPTA) for S and C-Band-Based Radar Applications

Murali Krishna Bonthu and Ashish Kumar Sharma

Abstract In this paper, the self-complimentary concept has been applied to Log Periodic Tripole Antenna (LPTA) to achieve the multiband resonance for S-band and C-band-based wireless applications. The proposed LPTA is designed with triangular shape elements in an array structure. This proposed self-complementary LPTA design exhibit nine multiband resonant frequencies at 2.50 GHz, 2.98 GHz, 3.31 GHz, 3.82 GHz, 4.54 GHz, 5.32 GHz, 6.31 GHz, 7.60 GHz and 8.95 GHz with corresponding bandwidths of 140 MHz, 190 MHz, 140 MHz, 220 MHz, 430 MHz, 400 MHz, 470 MHz, 680 MHz and 820 MHz, respectively, with adequate return loss > 17 dB. The proposed self-complementary LPTA shows the acceptable gain from 0.5 to 4.30 dB for S-band (2–4 GHz) and C-band (4–8 GHz) for surveillance and weather radar systems-based applications.

Keywords Log periodic diploe array (LPDA) · Microstrip patch · Return loss · Impedance bandwidth

1 Introduction

The self-complimentary antennas gained immense interest to exhibit broadband characteristics and multi band characteristics. The self-complimentary antennas were developed through the combination of Yagi-Uda and slot antennas [1–3]. The slot-based antenna deign show multiple resonant frequencies, which strongly rely on frequency dependent properties and constant input impedance of the antenna elements, which ultimately produce narrow bandwidth characteristics [2]. Therefore, the slot antenna gained immense attention as a low-profile antenna element. Later it was found that the introduction of the slots into a Yagi-Uda antenna shown tremendous performance improvement for Yagi-Uda antennas. Thereafter the planar conducting sheet with varying slot with identical to the shape of its complementary planar antenna termed as self-complementary antenna. These Self-complementary

M. K. Bonthu · A. K. Sharma (✉)
Department of Electronics and Telecommunications, Veer Surendra Sai University Technology, Burla, India
e-mail: ashishksharma29@gmail.com

247

antennas are based on the Mushiake relationship [1]. Therefore, log periodic antennas are considered to be derivatives of the self-complementary antennas, due to their constant input impedance independent of source frequency. Self-complimentary antennas are designed from Log Periodic Dipole Array (LPDA) to achieve wideband to a multiband resonant antenna [4, 5]. However, the LPDA designs show the field radiations on both the sides of the antenna element, which ultimately increase back side radiation. Hence self-complementary structures are used to minimize back side radiation with constant impedance for entire frequency band as compared to LPDA's. Hence the presence of ground plane with self-complimentary structure vanishes wideband behavior of the system and produces multiband Resonance [3, 6].

In this paper the self-complimentary tripole patch array designed to achieve multiband resonance in the range of S-band and C-band. The design seems like LPTA model with self-complimentary slots in the ground structure to exhibit multiband resonance in the UWB range for the S-band (2–4 GHz) and C-band (4–8 GHz) applications.

2   Design of Self-complimentary Log Periodic Tripole Antenna (LPTA)

This proposed self-complement LPTA has been designed on Flame Retardant (FR-4) epoxy glass material with dielectric constant of 4.4 and thickness of 1.6 mm with tangent of 0.02. Here the design process of self-complementary LPTA is similar to the Log Periodic Dipole Array (LPDA) antenna, except the dipole element is replaced with triangular shape element. Conventional LPDA geometry dimensions are calculated through Carrel design equations [7]. In this design the apex angle is defined as

$$\alpha = \tan^{-1}\left[\frac{1 - \tau}{4\sigma}\right] \tag{1}$$

Here LPTA has been designed with directivity of 8 dB, scaling factor (τ) of 0.865 and spacing factor (σ) of 0.157 for the frequency range 2–10 GHz. In this design the ratio of the lengths, width and spacing of two consecutive triangular elements is given as [8, 9]

$$\frac{1}{\tau} = \frac{a_{n+1}}{a_n} = \frac{b_{n+1}}{b_n} = \frac{s_{n+1}}{s_n} \tag{2}$$

The number of elements of an LPDA can be calculated as

$$N = 1 + \frac{\ln(B_s)}{\ln(1/\tau)} \tag{3}$$

(a) 3D View of Proposed Design

(b) Top View

(c) Bottom View
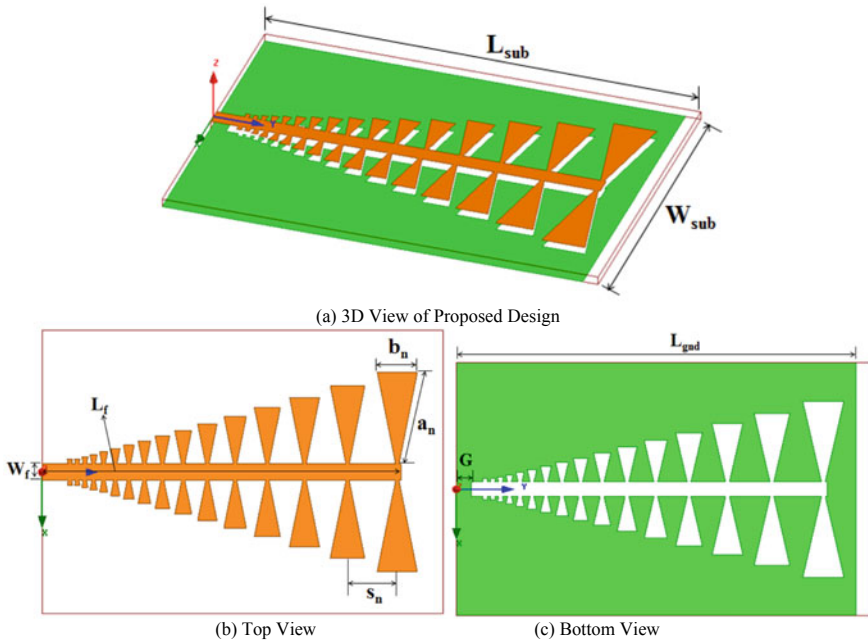
Fig. 1 Schematic of self-complementary log periodic tripole antenna (LPTA)

where the designed bandwidth (B_s) and desired bandwidth (B) is given as

$$B_s = B[1.1 + 7.7(1 - \tau)^2 \cot \alpha] \tag{4}$$

Here the total number of element $N = 16$ of the array is calculated through Eqs. (3) and (4). Hence sixteen triangle shaped elements are arranged in log periodic manner in order to improve bandwidth due to their irregular structure. Figure 1 shows the schematic of the proposed self-complementary LPTA with 50-Ω micro strip lines. This design represents a bilateral array of triangles on top of patch while the ground plane is designed as self-complimentary slots of bilateral array of triangles to achieve multiband resonance for S and C-band. The schematic of self-complementary LPTA is shown in Fig. 1 and all design parameters are given in Table 1.

3 Results and Discussion

This proposed multiband LPTA antenna has been simulated using Ansys High Frequency Structure Simulator (HFSS) software. Figure 2 shows the return loss characteristics of multiband LPTA design. This proposed Self-complementary LPTA obtains nine multiple bands resonance at 2.50, 2.98, 3.31, 3.82, 4.54, 5.32, 6.31, 7.60

Table 1 Design parameters for self-complementary LPTA (in mm)

L_{sub}	W_{sub}	L_f	W_f
85	60	76	3.5
L_{gnd}	G	h	
82	3.4	1.6	
Tripole (n)	a_n	b_n	S_n
1	21.54	8.47	10.536
2	18.63	7.33	9.113
3	16.11	6.34	7.883
4	13.94	5.48	6.819
5	12.05	4.74	5.898
6	10.43	4.10	5.102
7	9.02	2.65	4.413
8	7.80	3.07	3.817
9	6.75	2.65	3.302
10	5.84	2.29	2.856
11	5.05	1.98	2.470
12	4.37	1.72	2.137
13	3.78	1.49	1.848
14	3.27	1.28	1.599
15	2.83	1.11	1.383
16	2.42	0.96	1.196

and 8.95 GHz with corresponding bandwidths are 140, 190, 140, 220, 430, 400, 470, 680 and 820 MHz with corresponding return loss of − 30.17, − 20.98, − 19.97, − 31.91, − 22.79, − 23.42, − 22.36, − 31.16, − 35.02 dB for 2–10 GHz. Figure 3 shows the VSWR characteristics of proposed antenna deign and it is observed that the VSWR is < 2 for whole frequency band to achieve the impedance matching condition.

Table 2 shows the far filed plots of proposed antenna at all nine resonant frequencies. Here the far field plots show 3D gain polar plots and 2D radiation patterns. Here it is observed that the proposed self-complementary LPTA antenna design show the gain of 0.5 dB, 1.01 dB, 0.70 dB, 3.03 dB, 2.96 dB, 4.30 dB, 3.74 dB, 4.17 dB, 3.20 dB for resonant frequency of 2.50 GHz, 2.98 GHz, 3.31 GHz, 3.82 GHz, 4.54 GHz, 5.32 GHz, 6.31 GHz, 7.60 GHz and 8.95 GHz, respectively.

The resonance performance of proposed antenna design is represented through the surface current distributions of metallic triangular elements as shown in Fig. 4. Here it is observed that the maximum current is distributed in larger and smaller elements in order to show the resonance for 2–10 GHz, respectively. Here it is also found that the self-complementary triangular structure shows multiband resonance, which exhibit pre-filtering characteristics at the RF-front end [10].
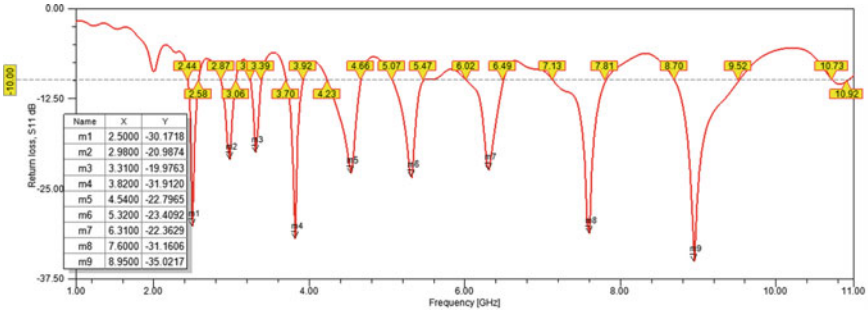
Fig. 2 S_{11} characteristics of self-complimentary slotted LPTA shows the multiple resonances
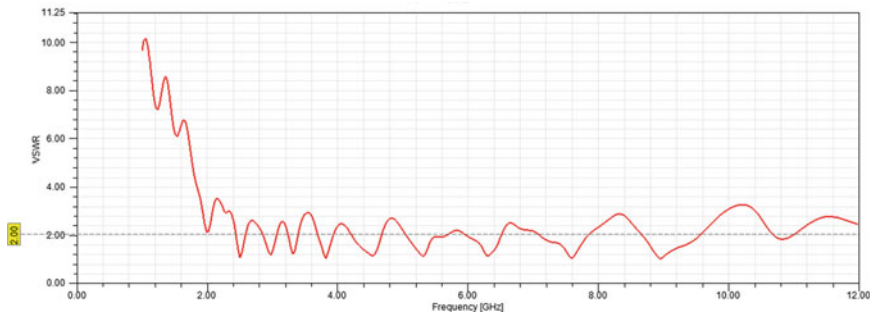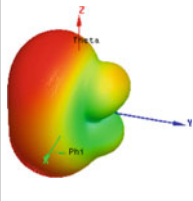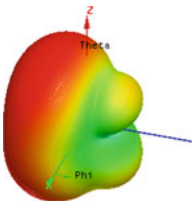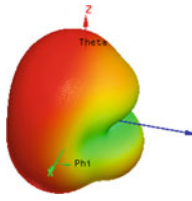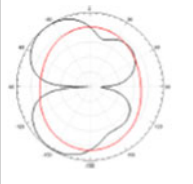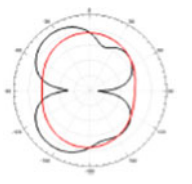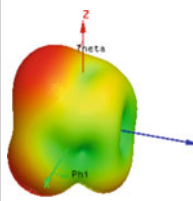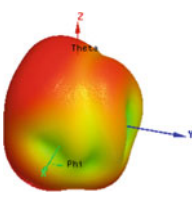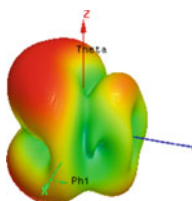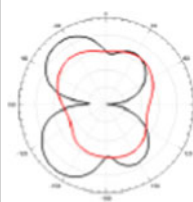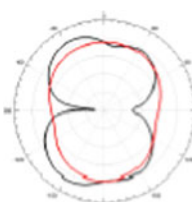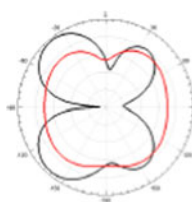
Fig. 3 VSWR characteristics of self-complimentary slotted LPTA

4 Conclusion
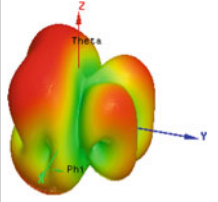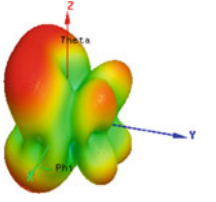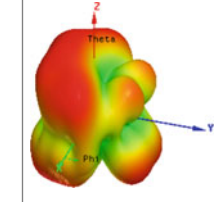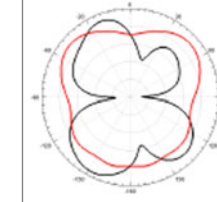
In this proposed design, it has been explained that the self-complimentary Log periodic tripole antennas suitable for multiband application in wide range of frequency from 2 to 10 GHz. The proposed self-complementary LPTA show the nine resonant frequencies at 2.50, 2.98, 3.31, 3.82, 4.54, 5.32, 6.31, 7.60 and 8.95 GHz with corresponding bandwidths of 140, 190, 140, 220, 430, 400, 470, 680 and 820 MHz. Hence the proposed design shows the major applications for WLAN/WiMAX, IEEE 802.11 and Bluetooth in S-band (2–4 GHz) and C-band (4–8 GHz) with adequate gain of 0.5, 1.01, 0.70, 3.03, 2.96, 4.30, 3.74, 4.17, 3.20 dB for nine resonant frequencies. Hence the self-complimentary LPTA show promising application for S-band (2–4 GHz) and C-band (4–8 GHz)-based surveillance and weather radar systems.

Table 2 Far filed reports of self-complimentary slotted LPTA (red—elevation plane and black—azimuth plane)

Resonant frequency (GHz)	2.50	2.98	3.31
Max. gain (dBi)	0.5	1.01	0.70
3D gain plot			
2D radiation pattern			
Resonant frequency (GHz)	3.82	4.54	5.32
Max. gain (dBi)	3.03	2.96	4.30
3D gain plot			
2D radiation pattern			

(continued)

Table 2 (continued)

Resonant frequency (GHz)	6.31	7.60	8.95
Max. gain (dBi)	3.74	4.17	3.20
3D gain plot			
2D radiation pattern			

(a) 2.50GHz

(b) 3.31GHz

(c) 6.31GHz

(d) 8.95GHz
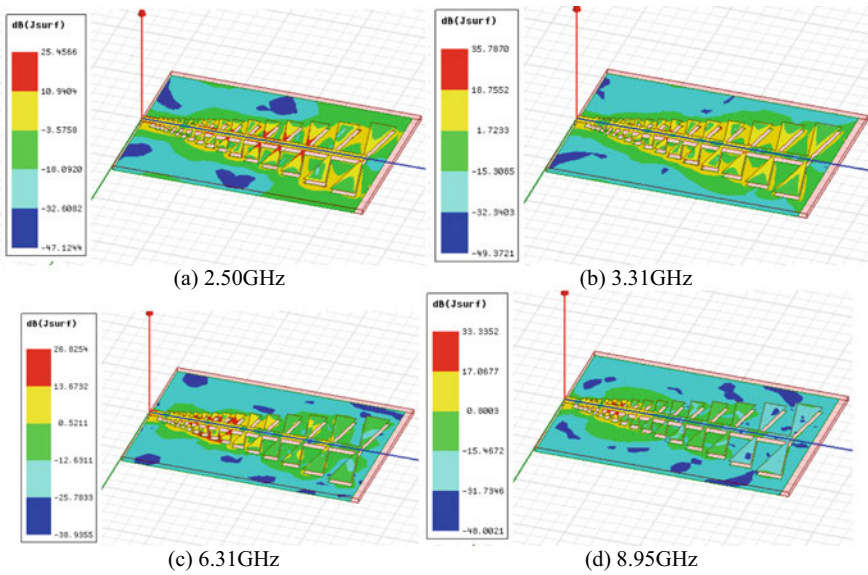
Fig. 4 Surface current distribution of self-complimentary slotted LPTA at some resonant frequencies

References

1. Mushiake Y (2004) A report on Japanese development of antennas: from the Yagi-Uda antenna to self-complementary antennas. IEEE Antennas Propag Mag 46(4):47–60. https://doi.org/10.1109/MAP.2004.1373999
2. Sawaya K, Ishizone T, Mushiake Y (2017) Principle of self-complementary and application to broadband antennas. In: IEEE history of electrotechnolgy conferences (HISTELCON). IEEE, pp 53–58. https://doi.org/10.1109/HISTELCON.2017.8535732
3. Dardenne X, Craeye C (2003) Simulation of the effects of a ground plane on the radiation characteristics of self-complementary arrays. In: IEEE antennas and propagation society international symposium. Digest. Held in conjunction with: USNC/CNC/URSI North American radio science meeting (Cat. No. 03CH37450), vol 1. IEEE, pp 383–386. https://doi.org/10.1109/APS.2003.1217477
4. Isbell DE (1960) Log periodic dipole arrays. IEEE Trans Antennas Propag 8:260–267
5. Casula GA, Maxia P, Mazzarella G, Montisci G (2013) Design of a printed log-periodic dipole array for ultra-wideband applications. Prog Electromagnet Res C 38:15–26. https://doi.org/10.2528/PIERC13012704
6. Abdo-Sanchez E, Esteban J, Martin-Guerrero TM, Camacho-Penalosa C, Hall PS (2014) A novel planar log-periodic array based on the wideband complementary strip-slot element. IEEE Trans Antennas Propag 62(11):5572–5580. https://doi.org/10.1109/TAP.2014.2357414
7. Carrel R (1966) The design of log-periodic dipole antennas. In: 1958 IRE international convention record, vol 9. IEEE, pp 61–75. https://doi.org/10.1109/IRECON.1961.1151016
8. Casula GA, Maxia P (2014) A multiband printed log-periodic dipole array for wireless communications. Int J Antennas Propag.https://doi.org/10.1155/2014/646394
9. Hall PS (1986) Multioctave bandwidth log-periodic microstrip antenna array. IEE Proc H (Microw Antennas Propag) 133(2):127–136. https://doi.org/10.1049/ip-h-2.1986.0021
10. Yu C, Hong W, Chiu L, Zhai G, Yu C, Qin W, Kuai Z (2010) Ultra wideband printed log-periodic dipole antenna with multiple notched bands. IEEE Trans Antennas Propag 59(3):725–732. https://doi.org/10.1109/TAP.2010.2103010

Performance of Cooperative Spectrum Sensing Techniques in Cognitive Radio Based on Machine Learning

S. Lakshmikantha Reddy and M. Meena

Abstract Due to the increasing demand for advanced wireless system applications such as long-range wireless power, vehicle-to-vehicle communication, various sensors' inter communication, using a lot of cloud data, wireless sensing, millimeter-wave wireless, software-defined radio (SDR), etc., more and more bandwidth is required which is possible by the spectrum sensing (SS) concept in cognitive radio (CR). Therefore, researchers are showing much interest in SS techniques along with machine learning (ML), since ML has been showing optimal solutions for various computational problems. Recently, ML became an emerging technology due to its advanced applications such as product recommendations, social media features, sentiment analysis, marine wildlife preservation, predict potential heart failure, language translation, automating employee access control, image recognition, regulating healthcare efficiency and medical services, banking domain, etc. In this paper, a detailed analysis of the most recent advances is presented about cooperative spectrum sensing (CSS), cognitive radio, and ML-based CSS.

Keywords Cognitive radio (CR) · Software-defined radio (SDR) · Machine learning (ML) · Cooperative spectrum sensing (CSS) · Spectrum sensing (SS)

1 Introduction

Recently, more and more multimedia applications are arising day by day that need a lot of spectrum. However, spectrum is a scarcity quantity and it is not possible to allot more spectrum for a single application. Therefore, in order to improve the spectrum efficiency, one of the most emerging technologies cognitive radio (CR) has been developed and its working modules are shown in Fig. 1 [1]. The static control

S. Lakshmikantha Reddy (✉) · M. Meena
Electronics and Communication Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, India
e-mail: slkreddy.phd@gmail.com

M. Meena
e-mail: meena.se@velsuniv.ac.in
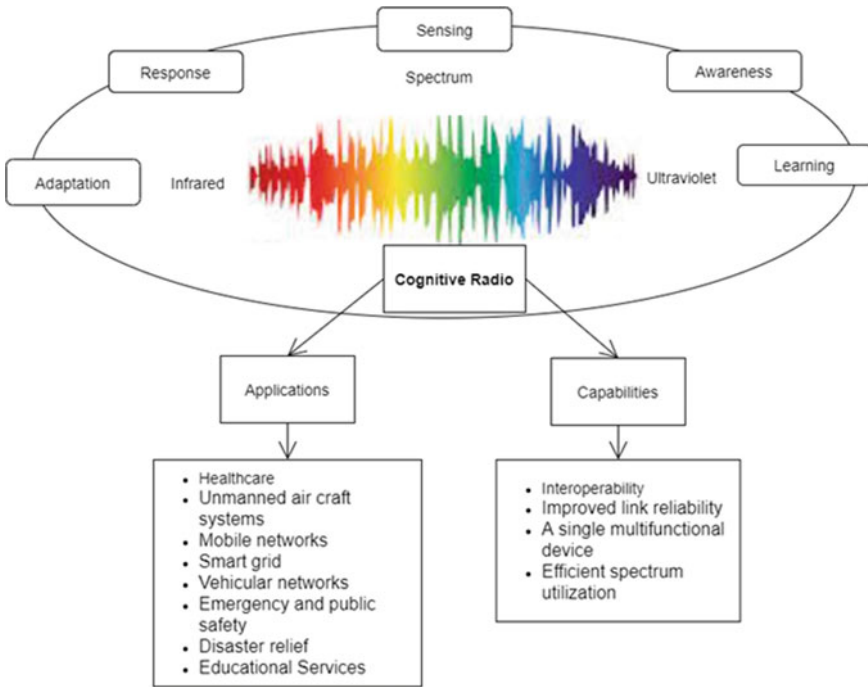
Fig. 1 Cognitive radio system with capabilities and applications

of the radio range is as of now not useful enough to give induction to this load of utilizations. Not splitting the radio range between clients can achieve the creation of bothersome withdrawal of administrations. The shortage of the radio range is consequently one of the most basic issues at the forefront of future research that actually just cannot be tended to. One solution for these and various troubles is to use cognitive radio innovation, which has gone through expansive assessment by the investigation neighborhood close to twenty years. Cognitive radio innovation might conceivably address the absence of available radio range by engaging unique range access. Since its show, analysts have been managing engaging this inventive advancement in managing the radio range. Therefore, this research field has been advancing at a quick speed and critical advances have been made [2–6]. Spectrum sensing is explained in detail in the upcoming sections.

1.1 Spectrum Sensing

It is the concept of allotting unlicensed user [secondary user (SU)] in the place of licensed users [primary user (PU)] when the PU is not using the spectrum. We found a number of sensing techniques in the literature (shown in Fig. 2) such as coordinated

with channel discovery [2], cyclostationary detection [3], covariance-based recognition [4], energy identification [5], and machine learning-based detection [6]. Current communication frameworks have execution prerequisites like enormous information volumes, high transmission rates, and quick reaction speeds, which pose difficulties to existing systems.

In view of the writing done [7], spectrum detecting is isolated into two distinct ways, for instance, cooperative mode and non-cooperative mode. In traditional CSS plans, a SU analyzes whether the PU's signal exists and makes a decision concerning the channel's available status. It is functional for the CR systems to take part to achieve higher distinguishing dependability quality than individual identifying does by yielding an unrivaled answer for the hidden PU issue that arises because of shadowing and multipath fading. While in cooperative detecting, the CR gadgets trade the recognizing results with the combination place for deciding. With hard combination calculations, the CR gadgets exchange a tiny bit of information with the combination place, which determines if the energy is over a particular breaking point. In the non-cooperative mode, each SU chooses the channel's status exclusively. This mode is fitting when the hubs cannot share their SS information. In this manner, ML-based ideas appeared to show the ideal answer for these sorts of hardships.
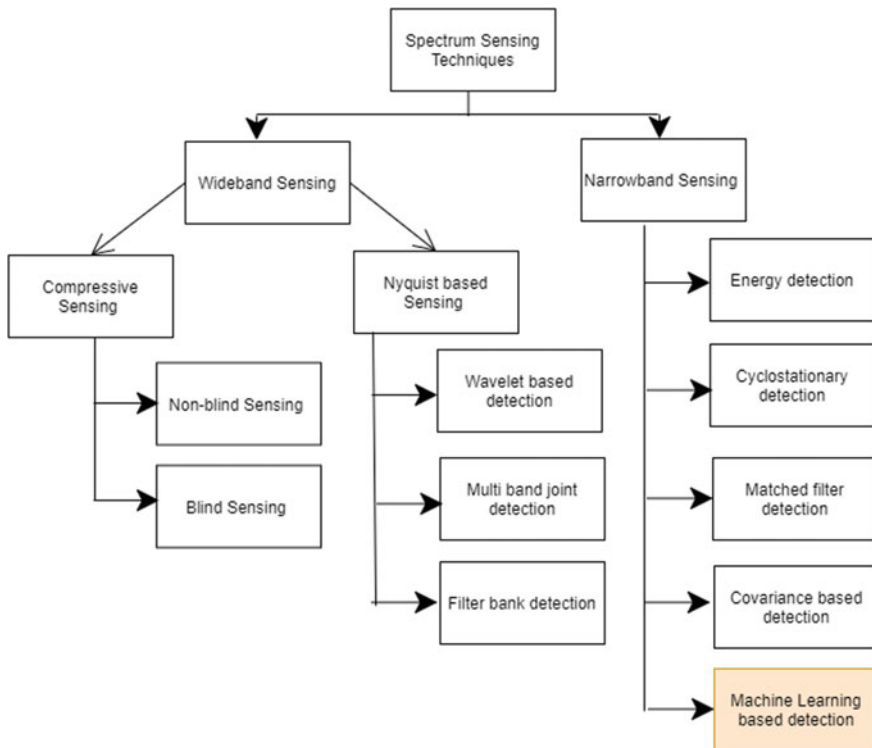
Fig. 2 Classification of spectrum sensing techniques

1.2 Machine Learning-Based SS

Generally, most of the machine learning techniques are developed for pattern classification, in which a feature vector is determined from the pattern with the help of feature classifiers and then the whole pattern is divided into a specific class. With regards to CSS, we treat an "energy vector", every part of which is an energy level assessed at every CR device, as a component vector. Then, at that point, the classifier sorts the energy vector into one of two classes: the "channel accessible class" (relating to the absence of PU) and the "channel inaccessible class" (comparing to the case that something like a minimum of one PU is present). Preceding on the web grouping, the classifier needs to go through a preparation stage where it gains from preparing highlight vectors. As per the kind of learning technique embraced, an algorithm is classified as unsupervised learning, supervised learning, and reinforcement algorithm as shown in Fig. 3. In supervised learning, a preparation highlight vector is taken care of into the classifier with its name demonstrating the real class the preparation include vector has a place with [8].

Bearing the current limitations in mind, Mi Y et al. [7] developed a technique to perform spectrum sensing for OFDM systems by taking the advantages of ML into consideration. The spectrum sensing issue is assumed as a multiple-class order issue, adjusting to obscure SNR varieties. It implies that the learning in our methodology is one-off, paying little heed to the next SNR varieties [8]. Thilina et al. [9], proposed a threshold technique to decide channel status is changed relying upon the PU's transmit power level. Liu et al. [10], proposed a DL-based CSS in which acknowledge signals are generated to represent the occupation of SUs in the place of PUs. In order to fix the scanning order of channels, Ning et al. [11], proposed a CSS method based on
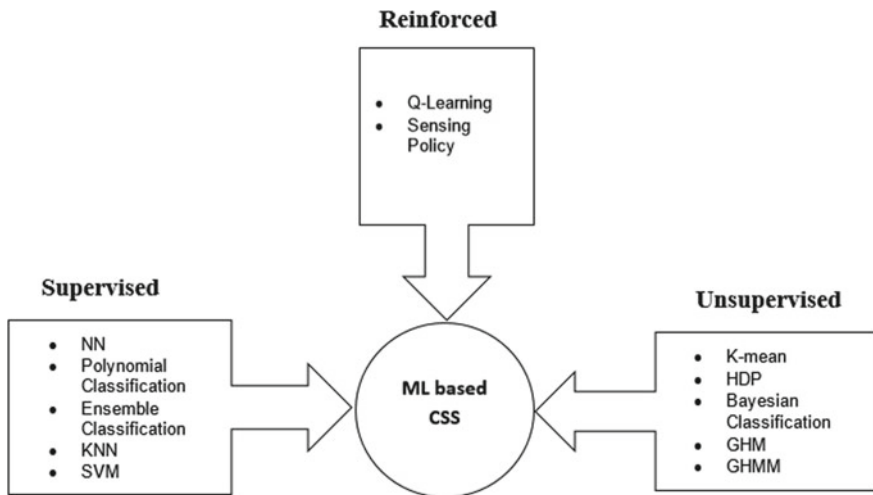
Fig. 3 Classification of machine learning

reinforcement learning (RL). Varun and Annadurai [12], proposed an ML-based CSS for medical applications. The proposed classifier utilizes the rule of high-speed need-based multi-layer utmost learning machines for the prediction and classification. The test testbed has planned based multicore CoxtexM-3 loads up for executing the constant intellectual situation and different execution boundaries like forecast exactness, preparing and testing time, receiver working qualities, and precision of discovery. Kim and Choi [13]., presented how to put sending nodes (SNs) to ensure the presentation of ML-based CSSs. Confirmed that the hidden PU issue makes the cross-over of information dissemination, which weakens the spectrum detecting capacity. In light of Kullback–Leibler divergence, insightful articulations for the spectrum detecting inclusion of a single SN are inferred. Proposed a technique on the best way to put a couple of SNs to cover the entire space of the PU and demonstrate the feasibility of the proposed method by experiment results.

He and Jiang [14], investigated the use of deep learning methods for wireless systems with attention to energy efficiency improvement for disseminated cooperative spectrum sensing. Fostered a deep learning structure by incorporating graph neural networks and support figuring out how to further develop the general system energy productivity. A successive sensor selection technique has been intended to track down an authentic subset of sensors that can satisfy the objective detecting execution, yet additionally, ensure the topology necessity of the distributed sensing algorithm. The authors in [15] analyzed exhaustively from the parts of the quantity of SUs, the preparation information volume, the normal signal-to-noise ratio of recipients, the proportion of PUs' power coefficients, just as the preparation time and test time. Multiple ML empowered solutions were acquainted to clear up the CSS issue. Broad mathematical outcomes showed that the proposed arrangements can accomplish precisely just as viable spectrum detecting and the DAG-SVM algorithm was more appropriate than different algorithms.

Wang and Liu [8], majorly analyzed cooperative schemes based on three ML algorithms such as convolutional neural network, support vector machine, and deep reinforcement learning a CSS strategy utilizing ELM and contrasted some famous ML procedures and some logical combination models. The proposed technique ends up being the most incredible in preparing time results than the wide range of various ML strategies. Besides, a correlation has been performed with some benchmark covariance-based techniques to demonstrate the strategy's comprehensiveness.

2 System Model

The cognitive radio system consists of four major components (shown in Fig. 4) such as the primary users, a base station for the primary user (P-BS), cognitive radio base station (CR-BS) also called as fusion center (FC), and cognitive radio users (secondary users). Primary users (PUs) are licensed (authorized) users who are having high priority to use the spectrum anytime, whereas secondary users (SUs) are non-licensed users. The spectrum is allotted to SUs in the absence of PUs only. The
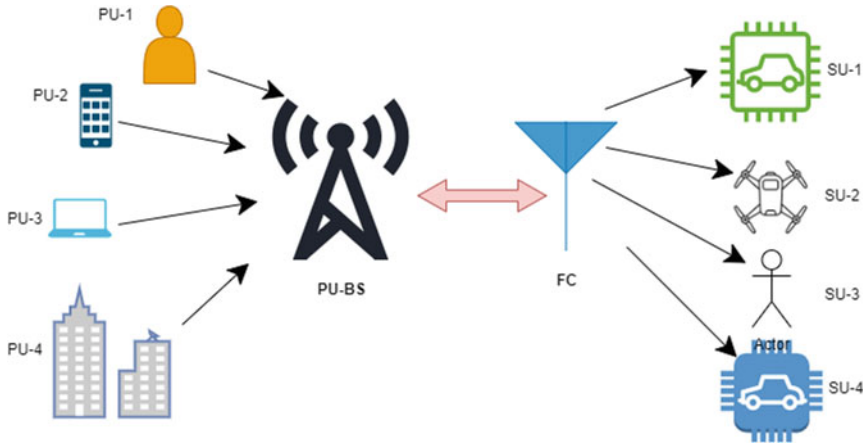
Fig. 4 Cooperative spectrum sensing model

spectrum allocation for PUs is done by the PU-BS and for SUs done by the CR-BS, however, CR-BS allotted spectrum to the SUs based on the priority.

The cooperative spectrum sensing working principle is almost the same as conventional spectrum sensing which is discussed at the beginning of this section. Whereas in CSS, a fusion center handles (in the place of CR-BTS) the whole spectrum sensing process by organizing sensors to report detecting information; combining the gathered detecting information, and making a decision whether the PU is present or absent on the spectrum.

When we try to depend on ML for CSS purposes, CSS can be described as a classification or testing problem. All the secondary users are represented in the form of vectors based on their energy, called an energy vector. The validation dataset provides an unbiased evaluation of a model fit on the training dataset while tuning the model's hyperparameters. All the elements of the vector are separated into training samples and test samples. Then, various ML algorithms can be used for classification purposes based on the application. The classifier presents the presence and absence of PU on each sub-band, and based on this information, free sub-bands are allocated to various SUs as per the order of priority.

3   Conclusion and Future Work

This paper presented a detailed survey on ML-based cooperative spectrum sensing techniques. In addition, the difference between conventional CSS and ML-based CSS is explained in detail along with proper figures. Further, a detailed literature survey on ML algorithms is presented by mentioning the proposed techniques with their significance and limitations. The researchers can focus more on advanced ML algorithms

to implement SS to achieve improved spectral efficiency. Further, can analyze and compare supervised, unsupervised, and reinforcement algorithms for CSS and non-CSS for multiple primary users. For practical proof of the concepts, the researchers can implement the proposed concepts on software-defined radio hardware.

References

1. Kaur A, Kumar K (2020) A comprehensive survey on machine learning approaches for dynamic spectrum access in cognitive radio networks. J Exp Theor Artif Intell 1–40
2. Ranjan A, Singh B (2016) Design and analysis of spectrum sensing in cognitive radio based on energy detection. In: Proceedings of the international conference on signal and information processing, Vishnupuri, India, pp 1–5
3. Alom MZ, Godder TK, Morshed MN, Maali A (2017) Enhanced spectrum sensing based on energy detection in cognitive radio network using adaptive threshold. In: Proceedings of the international conference on networking systems and security, Dhaka, Bangladesh, pp 138–143
4. Arjoune Y, El Mrabet Z, El Ghazi H, Tamtaoui A (2018) Spectrum sensing: enhanced energy detection technique based on noise measurement. In: Proceedings of the IEEE computing and communication workshop and conference (CCWC), Las Vegas, NV, USA, pp 828–834
5. Chen A-Z, Shi Z-P (2020) Covariance-based spectrum sensing for noncircular signal in cognitive radio networks with uncalibrated multiple antennas. IEEE Wireless Commun Lett 9(5):662–665
6. Yang T et al (2019) Fusion rule based on dynamic grouping for cooperative spectrum sensing in cognitive radio. IEEE Access 7:51630–51639
7. Mi Y et al (2019) A novel semi-soft decision scheme for cooperative spectrum sensing in cognitive radio networks. Sensors 19(11):2522
8. Wang J, Liu B (2021) A brief review of machine learning algorithms for cooperative spectrum sensing. J Phys: Conf Ser 1852(4)
9. Thilina KM, Choi KW, Saquib N, Hossain E (2013) Machine learning techniques for cooperative spectrum sensing in cognitive radio networks. IEEE J Sel Areas Commun 31(11):2209–2221
10. Liu S, He J, Wu J (2021) Dynamic cooperative spectrum sensing based on deep multi-user reinforcement learning. Appl Sci 11(4):1884
11. Ning W et al (2020) Reinforcement learning enabled cooperative spectrum sensing in cognitive radio networks. J Commun Netw 22(1):12–22
12. Varun M, Annadurai C (2021) PALM-CSS: a high accuracy and intelligent machine learning based cooperative spectrum sensing methodology in cognitive health care networks. J Ambient Intell Humaniz Comput 12(5):4631–4642
13. Kim J, Choi JP (2019) Sensing coverage-based cooperative spectrum detection in cognitive radio networks. IEEE Sens J 19(13):5325–5332
14. He H, Jiang H (2019) Deep learning based energy efficiency optimization for distributed cooperative spectrum sensing. IEEE Wirel Commun 26(3):32–39
15. Shi Z et al (2020) Machine learning-enabled cooperative spectrum sensing for non-orthogonal multiple access. IEEE Trans Wireless Commun 19(9):5692–5702

Unified Power Quality Conditioner for V-V Connected Electrified Railway Traction Power Supply System

Ruma Sinha, H. A. Vidya, and H. R. Sudarshan Reddy

Abstract AC traction load is generally a large single-phase load with dynamic characteristics due to which a significant negative sequence component of current (NSC) is drawn from the grid and harmonics are injected in voltage and current. This inflicts undesirable effects on power quality. This work proposes a unified power quality conditioner which is a combination of a shunt-connected distribution static synchronous compensator (D-STATCOM) and a series-connected dynamic voltage restorer (DVR) to mitigate the degradation of power quality. The D-STATCOM helps in reducing the NSC significantly and suppressing the harmonic distortions. DVR addresses the voltage sag caused by the shunt compensating device. Control strategies for D-STATCOM and DVR for the unified power quality compensator are presented. The system is simulated in MATLAB environment, and a comparative analysis is carried out under various loading circumstances. As a result, superior power quality is achieved with the proposed UPQC.

Keywords Railway traction system · Unified power quality conditioner · Power quality · Negative sequence component of current · NSC · UPQC

1 Introduction

The popularity of electrified railway transportation is blooming exponentially due to multi-fold advantages such as decreased dependency on fossil fuel, de-carbonization, cleanliness, safety, etc. Mainline traction in India is deployed with 25 kV, 50 Hz AC power supply through overhead catenary. In the receiving substation, three-phase AC

R. Sinha (✉)
Global Academy of Technology, Bangalore, Karnataka, India
e-mail: ruma.sinha@gat.ac.in

H. A. Vidya
Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: ha_vidya@blr.amrita.edu

H. R. Sudarshan Reddy
University B.D.T. College of Engineering, Davangere, Karnataka, India

is converted to two single-phase supplies using a suitable transformer connection. The load in these two phases is seldom balanced, which leads to unbalanced current dawn from the power grid. Thus, the upstream power supply gets impacted with negative sequence component of current (NSC), total harmonic distortion (THD), voltage sag/swell, and other power quality disturbances [1–4]. Nonlinear and dynamic characteristics of traction load in addition to the large-scale use of power electronic converters employed in various stages for power conversion also causes power quality (PQ) issues [5]. Power quality disturbances affect the upstream power supply and are a cause of concern for the other consumers connected to the system. These PQ issues may cause malfunctioning or false triggering of protective relays and may affect railroad signaling and communication system [6]. Compensation of PQ issues is, therefore, a real necessity.

Researchers have proposed a variety of technologies in recent decades to alleviate the power quality issues [7–11]. Passive filters are a common solution but can provide compensation only for fixed order harmonics and are unable to provide dynamic compensation. C type passive power filter is widely used to filter out specific order of harmonic and to inject reactive power. But they can only inject fixed VAR which may lead to overcompensation and power factor degradation [8]. Poor flexibility of passive filters can be overcome using active filters. These active filters have been tested in a variety of configurations to improve power quality [12–15]. Some specially connected balanced transformers are proposed for the compensation but since the load in two phases is rarely balanced, these transformers are unable to compensate the NSC [9, 10]. A "static railway power conditioner" (RPC) is proposed, which is capable to provide overall compensation, requires reduced floor space and has flexible control [11]. However, Railway power conditioner (RPC) is not able to eliminate the neutral zones. Co-phase systems are able to eliminate neutral zones. A co-phase traction power supply system is proposed to solve PQ issues and eliminate neutral sections in the substation [16]. But complete elimination of neutral section may not be possible. Using multilevel inverter topology and pulse width modulation techniques, issues related to harmonics can be reduced [17, 18]. However, this calls for a requirement of higher traction power which in turn is responsible to draw more NSC from the three-phase supply. In order to eliminate neutral section completely, a multilevel converter-based system is proposed [19]. But complexity increases with the increase in number of levels. Unified power conditioner has been presented in the literature to mitigate power quality issues [20–22]. However, with the knowledge of authors, it was felt that application of unified power quality conditioner in the field of railway power quality conditioning has not been carried out in detail. With this motivation, the paper proposes a unified power quality conditioner to mitigate the power quality issues with respect to harmonic distortion, the negative sequence component of current, and voltage sag in the upstream power supply network. The system performance is compared with an installation of D-STATCOM and a DVR separately.

1.1 Traction System Configuration

For the single power supply to overhead catenary from three-phase power, various transformer connections can be used such as V-V, Scott, Leblanc, etc. [9, 10]. Among these, V-V connection has a simple topology and also has a good power utilization ratio. Thus, a V-V connected transformer is adopted for this work. The connection diagram and vector diagram are depicted in Fig. 1. V_M and V_T represent the main phase and teaser phase voltage. The traction load is connected in the main phase, and the auxiliary load is connected in the teaser phase. Due to the dynamic characteristic of traction load, a negative sequence current is drawn from the supply which may be as high as positive sequence component.

25 kV from the overhead catenary is collected using pantographs. WAP-7 locomotive uses three-phase squirrel-cage induction motors of 850 kW (1140 HP), 2180 V, 1283/2484 rpm, 270/310 A, and 6330 N-m [23]. In addition to the traction motors, auxiliary loads are present for which a single-phase power supply is required. In this paper, the load is simulated considering the three-phase traction motor load and single-phase auxiliary loads.

2 Design of D-STATCOM to Improve the Power Quality of Traction System

Voltage source converter (VSC)-based D-STATCOM is designed and connected at the distribution side as depicted in Fig. 2 to improve the power quality [24, 25]. The voltage at the DC bus is required to be at least two times higher than the peak voltage of the system and is obtained as

$$V_{DC} = \frac{2\sqrt{2}V_{LL}}{\sqrt{3}m} \tag{1}$$

(a) (b)
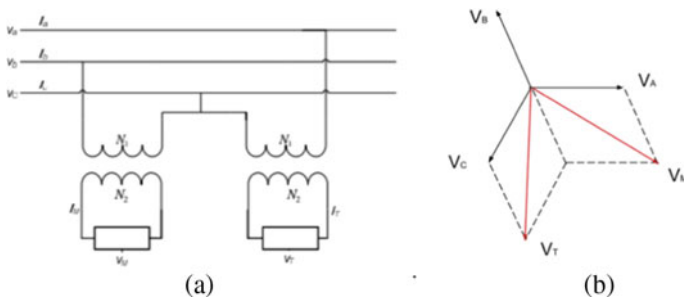
Fig. 1 V-V connected transformer **a** connection diagram **b** vector diagram

Fig. 2 System structure
with D-STATCOM
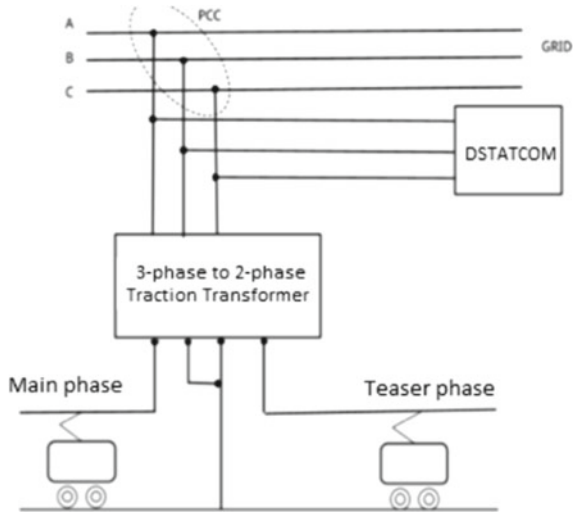
where m = modulation index and chosen as 1. A capacitor is connected to maintain the
DC link voltage. This capacitor is responsible to supply the energy during transient
condition and is calculated from the equation given as

$$\frac{1}{2}C_{DC}\left(V_{DC}^2 - V_{DCm}^2\right) = 3K_1 V_{ph} a I_{ph} t \tag{2}$$

where V_{DC} = reference DC voltage, a = overloading factor, t = desired recovery
time and K_1 is allowable energy imbalance during transients, and is considered as
0.1. Inductance connected in each leg of VSC is calculated as

$$L_f = \frac{\sqrt{3}m V_{DC}}{12 a f_s I_{ripple}} \tag{3}$$

where f_s is the switching frequency.

2.1 Control Algorithm of D-STATCOM

The load currents are initially sensed in an a-b-c frame of reference and then converted
to a d-q-0 frame of reference using Eq. 4.

$$i_{Lq} = \frac{2}{3}i_{La} \cos \omega t + \frac{2}{3}i_{Lb} \cos\left(\omega t - \frac{2\pi}{3}\right) + \frac{2}{3}i_{Lc} \cos\left(\omega t + \frac{2\pi}{3}\right)$$

$$i_{Ld} = \frac{2}{3}i_{La} \sin \omega t + \frac{2}{3}i_{Lb} \sin\left(\omega t - \frac{2\pi}{3}\right) + \frac{2}{3}i_{Lc} \sin\left(\omega t + \frac{2\pi}{3}\right)$$

$$i_{Lo} = \frac{1}{3}i_{La} + \frac{1}{3}i_{Lb} + \frac{1}{3}i_{Lc} \tag{4}$$

where

$\sin\omega t$ and $\cos\omega t$ are calculated by using a three-phase phase locked loop (PLL) over point of common coupling (PCC) voltage. The current thus obtained includes both DC component as well as ripple component. To extract the DC component, a low pass filter is used.

$$i_{Ld} = i_{dDC} + i_{dAC}$$
$$i_{Lq} = i_{qDC} + i_{qAC} \tag{5}$$

The obtained *d-q-0* current is converted into *a-b-c* reference frame using Park's transformation to generate the reference signal for the PWM pulse generation using Eq. 6 (Fig 3).

$$i_{sa}^* = i_q^* \cos(\omega t) + i_d^* \sin(\omega t) + i_o^*$$
$$i_{sb}^* = i_q^* \cos\left(\omega t - \frac{2\pi}{3}\right) + i_d^* \sin\left(\omega t - \frac{2\pi}{3}\right) + i_o^*$$
$$i_{sc}^* = i_q^* \cos\left(\omega t + \frac{2\pi}{3}\right) + i_d^* \sin\left(\omega t + \frac{2\pi}{3}\right) + i_o^* \tag{6}$$

Implementation of D-STATCOM in traction power supply system improves the harmonic distortion and works well for NSC compensation, but this affects the voltage profile and results in voltage sag.
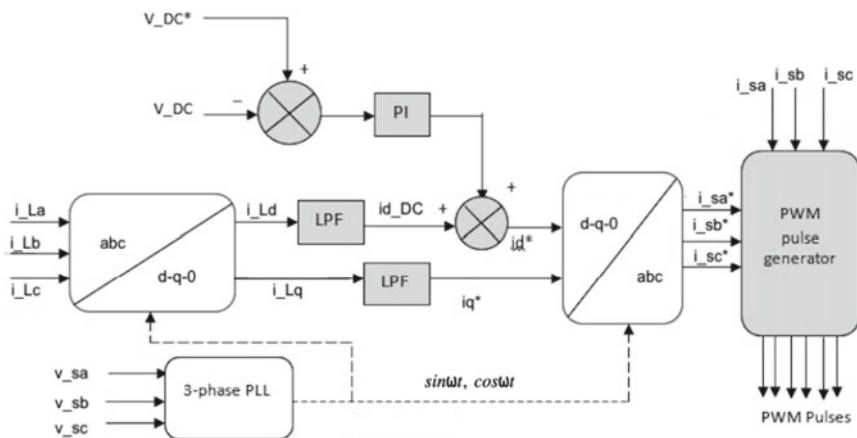
Fig. 3 Pulse generation for D-STATCOM

3   Design of UPQC Controller to Improve the Power Quality of Traction System

Unified power quality compensator is designed which is a combination of a D-STATCOM as shunt compensator and a DVR as series compensator (Fig. 4). D-STATCOM reduces harmonic distortion and NSC, while the DVR compensates the voltage sag. The capacitor helps in the voltage stabilization of the DC link.

3.1   Design of DVR

DVR is designed to compensate the voltage variations. To compensate a variation of $\pm X\%$, voltage to be injected by DVR is calculated as

$$V_{\text{DVR}} = x V_{\text{s}} \tag{7}$$

Turns ratio of the injection transformer is obtained as

$$K_{\text{DVR}} = \frac{V_{\text{s}}}{\sqrt{3}V_{\text{DVR}}} \tag{8}$$

The filter inductance of the DVR is obtained using Eq. 9.

$$L_{\text{DVR}} = \left(\frac{\sqrt{3}}{2}\right) m_a V_{\text{DC}} K_{\text{DVR}}/6a f_s I_{\text{DVR}} \tag{9}$$

where I_{DVR} is the current ripple in the supply current. A high-pass first-order filter is employed to filter out the higher frequency noise and tuned at 50% of the switching frequency.

3.2   Control Strategy of DVR

The control strategy of DVR is illustrated in Fig. 5. Synchronous reference frame theory is used for the generation of reference signal. The voltage at PCC in *a-b-c* reference frame is first converted to *d-q*-0 reference frame using "Park's transformation" as shown in Eq. 10.

$$v_{sq} = \frac{2}{3}v_{sa}\cos\omega t + \frac{2}{3}v_{sb}\cos\left(\omega t - \frac{2\pi}{3}\right) + \frac{2}{3}v_{sc}\cos\left(\omega t + \frac{2\pi}{3}\right)$$

$$v_{sd} = \frac{2}{3}v_{sa}\sin\omega t + \frac{2}{3}v_{sb}\sin\left(\omega t - \frac{2\pi}{3}\right) + \frac{2}{3}v_{sc}\sin\left(\omega t + \frac{2\pi}{3}\right)$$

Fig. 4 General structure of UPQC
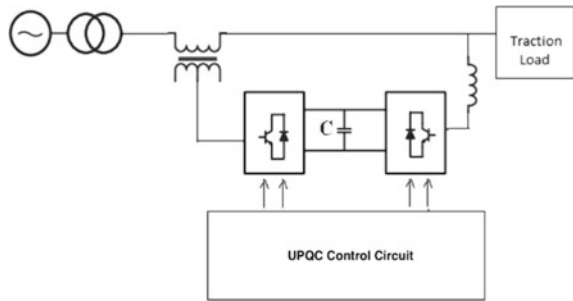
$$v_{so} = \frac{1}{3}v_{sa} + \frac{1}{3}v_{sb} + \frac{1}{3}v_{sc} \tag{10}$$

where $\sin\omega t$ and $\cos\omega t$ are calculated by using a three-phase PLL over grid side current.

The voltage thus obtained includes both DC component as well as ripple component. To extract the DC component, a low pass filter is used.

$$v_{sd} = v_{d\text{DC}} + v_{Ld\text{AC}}$$
$$v_{sq} = v_{q\text{DC}} - v_{Lq\text{AC}} \tag{11}$$

The magnitude of load voltage is obtained as

$$V_L = \sqrt{\frac{2}{3}}\left(\sqrt{v_{La}^2 + v_{La}^2 + v_{Lc}^2}\right) \tag{12}$$

Voltage regulation is obtained using proportional plus integral controller. The d and q axis reference voltages are estimated by using Eq. 13.

$$v_{Lq}^* = v_{q\text{DC}} + v_{qr}$$
$$v_{Ld}^* = v_{d\text{DC}}$$
$$v_{L0}^* = 0 \tag{13}$$

4 Simulation Result and Discussion

In order to analyze the system performance, a traction power supply system using V-V transformer is simulated in MATLAB Simulink environment. To model the traction system load, single-phase auxiliary load is also included in addition to the actual traction motor load. Simulation model of overall system and the traction load are shown in Figs. 6 and 7, respectively. System performance is analyzed with variation
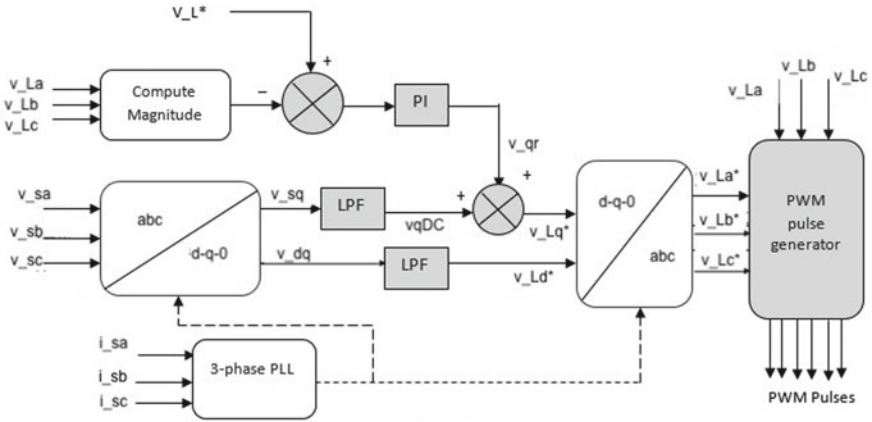
Fig. 5 Pulse generation of DVR

of load balance degree under the following stages: (A) without any compensator, (B) with D-STATCOM, (C) with DVR, and (D) with UPQC.
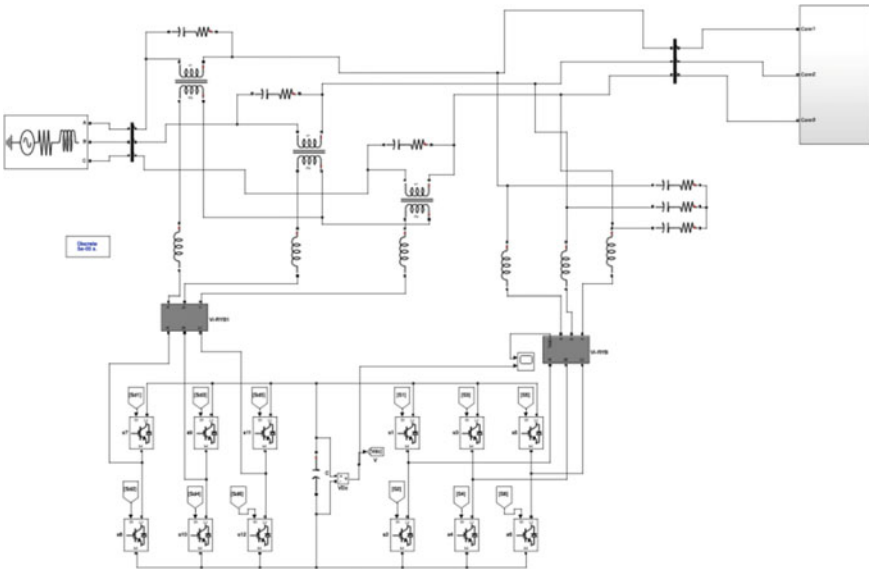
Fig. 6 MATLAB Simulink model of UPQC
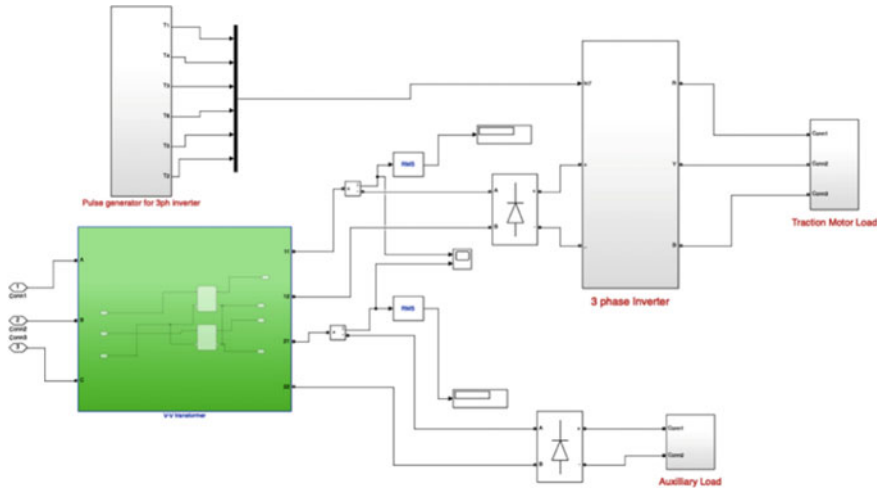
Fig. 7 Modeling of traction load

4.1 Without Any Compensator

First, the system performance is investigated without connecting any compensators. The source voltage and current suffers from harmonic distortion and system unbalance. In traction system, the load connected in the two phases is seldom balanced. Due to the unbalance in the current, a NSC is drawn which increases with the unbalance in the load. Grid side voltage and current waveform, total harmonic distortion, and the PSC, NSC, and ZSC are depicted in Fig. 8 as obtained from simulation.

The load connected in the main phase and the teaser phase is not balanced. Load balance degree is defined as the current drawn by the low load side to the current drawn by the high load side. It is observed that even when the load is completely balanced, i.e., 100% load balance degree, the load draws NSC from the source. When the load balance degree reduces, the negative sequence component of current drawn from the supply increases.

4.2 With D-STATCOM

A voltage source inverter-based D-STATCOM is designed and connected at the grid side. The performance is studied for different load balance degrees in the main and teaser phase. The performance is studied for different load balance degrees in the main and teaser phase. The graphs as obtained from simulation and the harmonic distortions in the source side voltage and current are illustrated in Fig. 9. The study shows with the appropriate design of D-STATCOM, harmonic distortion in grid side voltage, and current can be reduced. A significant reduction is achieved in the NSC

(a) Grid Side Voltage and Current

(b) Positive, Negative and Zero Sequence Component of Current

(c) Grid Voltage THD

(d) Grid Current THD
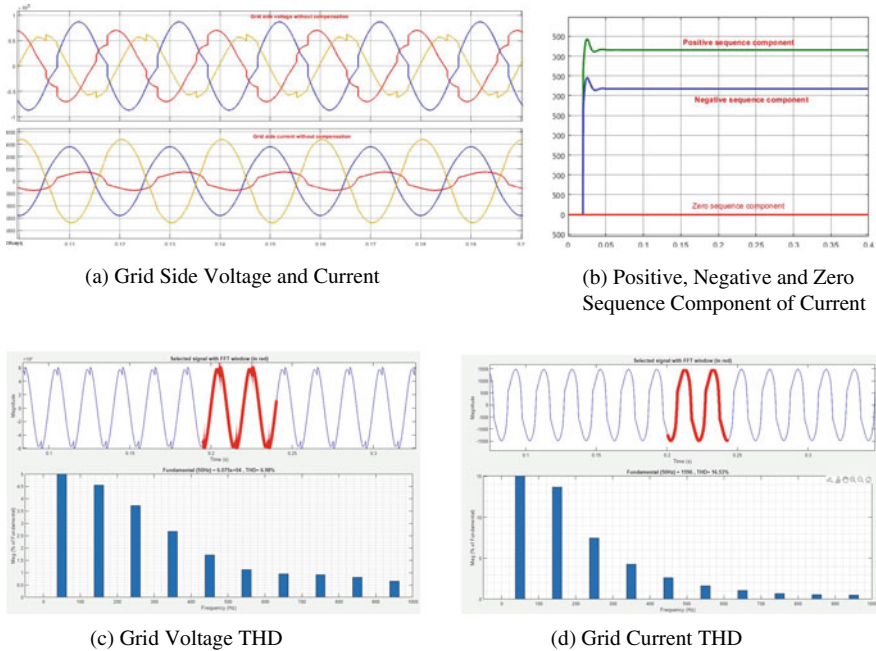
Fig. 8 Impact of traction load on grid side voltage and current without any compensator connected

drawn from the supply as well. However, it has been observed that the connection of D-STATCOM, in this case, causes a significant voltage sag in the source side, which would affect the other consumers connected to the utility.

4.3 With DVR

To compensate this voltage sag, a series compensator DVR is connected to obtain unified power quality compensation. VSC-based DVR injects a voltage through injection transformer and restores the required voltage. DVR alone can restore the voltage but cannot mitigate the problem of harmonic distortions. The results obtained with connection of a DVR as illustrated in Fig. 10. Finally, unified power quality compensator is connected which is a combination of a shunt compensator, D-STATCOM, and a series compensator, DVR.

(a) Grid Side Voltage and Current

(b) Positive, Negative and Zero Sequence Component of Current

(c) Grid Voltage THD

(d) Grid Current THD
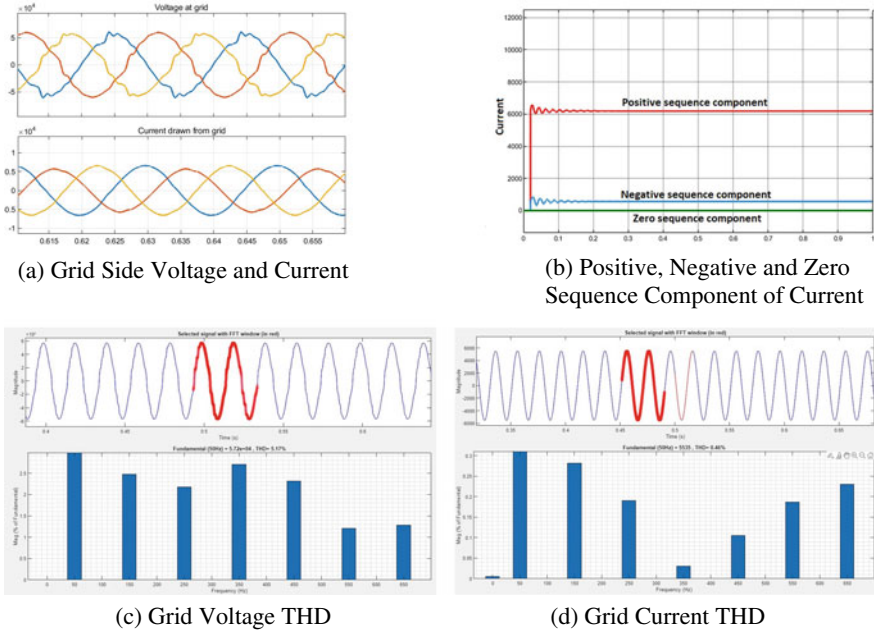
Fig. 9 Impact of traction load on grid side voltage and current with D-STATCOM connected

(a) Grid Side Voltage and Current

(b) Positive, Negative and Zero Sequence Component of Current

(c) Grid Voltage THD

(d) Grid Current THD
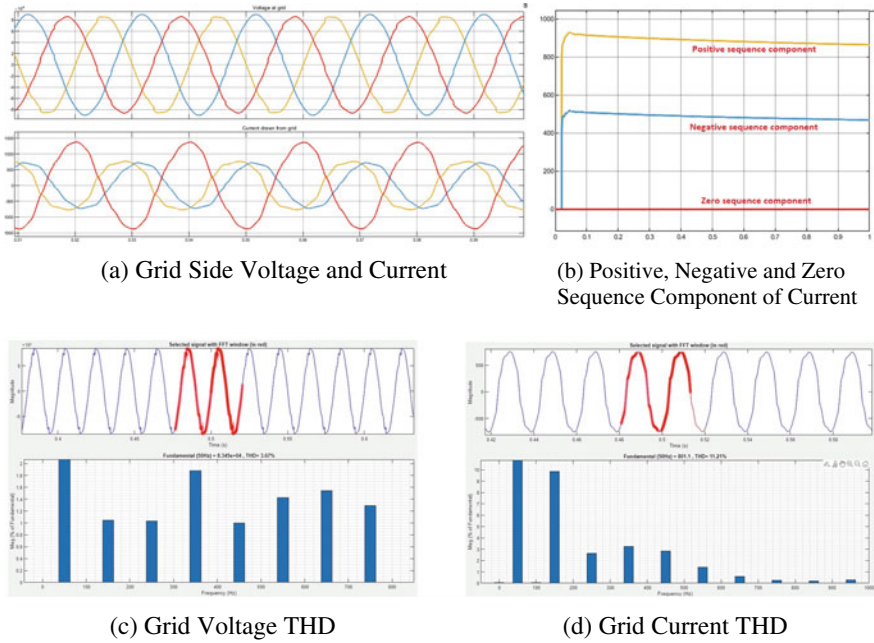
Fig. 10 Impact of traction load on grid side voltage and current with DVR connected

4.4   With UPQC

The total harmonic distortion obtained in grid side voltage and current obtained with
UPQC are almost negligible. Current unbalance and voltage sag also gets minimized
with the implemented VSC as shown in Fig. 11. Three goals have been achieved by
utilizing this UPQC in traction system. (1) Significant reduction of NSC drawn from
the grid, (2) suppression of harmonic distortion in grid side voltage and current, and
(3) limiting the voltage sag.

The system parameters as calculated using Eqs. 1–3 and 7–9 are given in Table 1.

(a) Grid Side Voltage and Current

(b) Positive, Negative and Zero
Sequence Component of Current

(c) Grid Voltage THD

(d) Grid Current THD
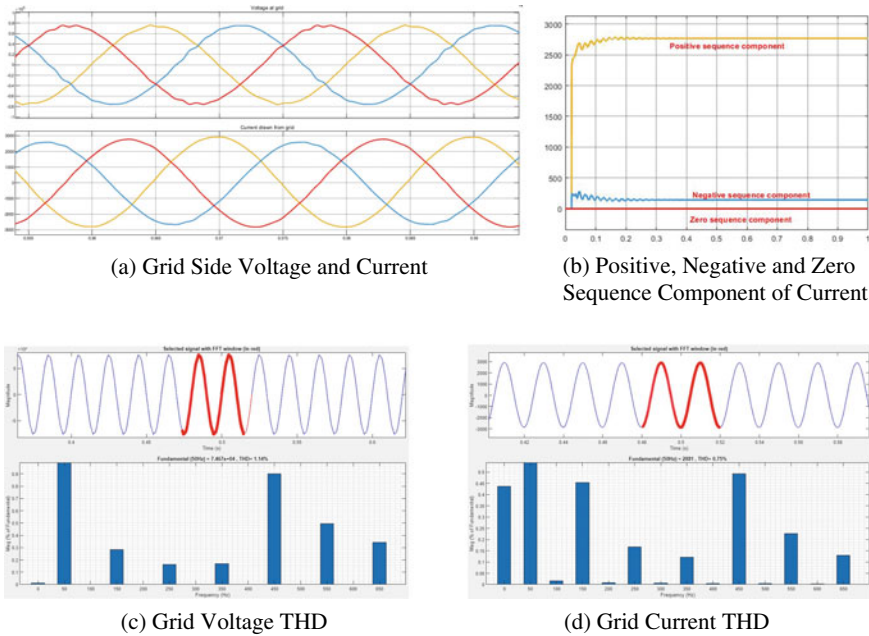
Fig. 11 Impact of traction load on grid side voltage and current with UPQC connected

Table 1 System parameters

System parameter	Value
Voltage at receiving station	11,000 V
DC bus voltage, V_{DC}	18,100 V
DC link capacitor, C_{DC}	2.4 mF
Filter inductance of D-STATCOM, L_f	33 mH
Turns ratio of injection transformer, K_{DVR}	2
Filter inductance of DVR, L_{DVR}	146 mH

(a) Variation of Voltage Sag with Load Balance Degree

(b) Current Unbalance with Variation of Load Balance Degree
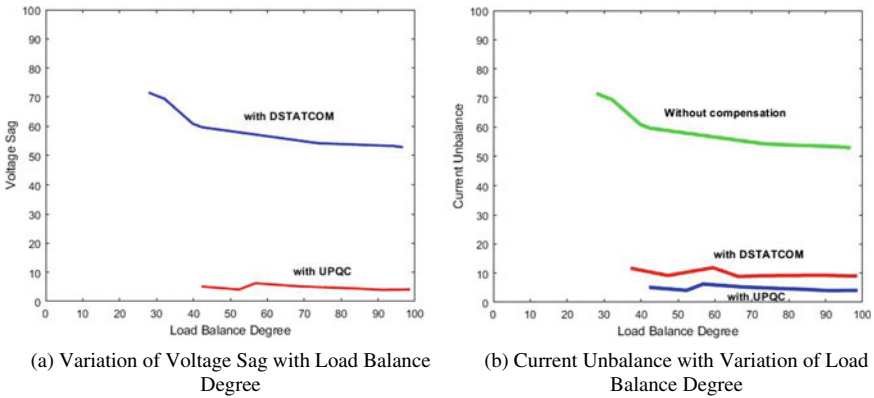
Fig. 12 Current unbalance and voltage sag variation with load balance degree

Table 2 Impact on source side voltage THD, current THD, current unbalance, and voltage sag with the variation of load balance degree and compensator connected

System	Source side voltage THD (%)	Source side current THD (%)	Load balance degree (%)	Current unbalance (%)	Voltage sag (%)
Without Compensator	6.97 4.95	16.53 9.39	50 88	55.4 48	– –
With D-STATCOM	5.15 3.75	0.43 0.34	45 82	20.2 19.7	56.5 56.8
With DVR	3.61	11.21	72.6	46.6	1.6 (swell)
With UPQC	0.85 1.18	0.3 0.8	46.7 84.9	0.6 4.4	9.4 8.6

With the variation of load balance degree, the variation of current unbalance, and the voltage sag is illustrated in Fig. 12 for uncompensated system, system with D-STATCOM, and UPQC (Table 2).

5 Conclusion

The dynamic characteristics of single-phase heavy-duty traction load impacts the grid voltage and current with harmonic distortions draws NSC and causes other power quality issues. As the traction power supply involves a three-phase to two one-phase conversion, the unbalance in grid side current increases with the decrease in load balance degree. This paper proposes a unified power quality compensator to improve the power quality of V-V connected electrified railway traction system.

The result is compared with a D-STATCOM and a DVR. The implementation of D-STATCOM mitigates the issues related to the harmonic distortions and NSC drawn from supply; however, this leads to voltage sag.

A dynamic voltage restorer as a series compensator is designed to indemnify the voltage sag. However, DVR alone is not capable enough to mitigate harmonic distortion and NSC. A combination of D-STATCOM, as a shunt device, and a DVR as a series compensator is connected for this purpose to obtain unified compensation. A detailed study reveals that with the UPQC the supply side voltage and current THD can be brought down to the permissible limit, NSC drawn from the supply is brought down to a very low value and voltage sag also is reduced. Performance analysis is carried out under different load balance degree, and it is observed that superior power quality is achieved with UPQC under both balanced as well as unbalanced load conditions.

References

1. Gazafrudi SMM, Langerudy AT, Fuchs EF, Al-Haddad K (2015) Power quality issues in railway electrification: a comprehensive perspective. IEEE Trans Ind Electron 62:3081–3090
2. Chen T-H, Yang W-C, Hsu Y-F (1998) A systematic approach to evaluate the overall impact of the electric traction demands of a high-speed railroad on a power system. IEEE Trans Veh Technol 47(4)
3. Kaleybar HJ, Brenna M, Foiadelli F, Fazel SS, Zaninelli D (2020) Power quality phenomena in electric railway power supply systems: an exhaustive framework and classification. Energies 13:6662
4. Morrison RE (2000) Power quality issues on AC traction systems. In: Proceedings of the Ninth international conference on harmonics and quality of power (Cat. No.00EX441), , vol 2, pp 709–714. https://doi.org/10.1109/ICHQP.2000.897765
5. Hu H, Shao Y, Tang L, Ma J, He Z, Gao S (2018) Overview of harmonic and resonance in railway electrification systems. IEEE Trans Ind Appl 54:5227–5245
6. Foley FJ (2011) The impact of electrification on railway signalling systems. In: IET professional development course on railway electrification infrastructure and systems (REIS), pp 146–153
7. Chen Y, Chen M, Tian Z, Liu L, Hillmansen S (2019) VU limit pre-assessment for high-speed railway considering a grid connection scheme. IET Gener Transm Distrib 13:1121–1131
8. Lamlom A, Ibrahim A, Balci ME, Karadeniz A, Aleem SHA (2017) Optimal design and analysis of anti-resonance C-type high-pass filters. In: Proceedings of the 2017 IEEE international conference on environment and electrical engineering and 2017 IEEE industrial and commercial power systems europe (EEEIC/I & CPS Europe), Milan, Italy, pp 1–6
9. Ciccarelli F, Fantauzzi M, Lauria D, Rizzo R (2012) Special transformers arrangement for AC railway systems. In: Electrical systems for aircraft, railway and ship propulsion (ESARS), pp 1–6
10. Vasanthi V, Ashok S (2011) Harmonic filter for electric traction system. In: IEEE PES innovative smart grid technologies—India
11. Dai N, Lao K, Wong M (2013) A hybrid railway power conditioner for traction power supply system. In: Twenty-eighth annual IEEE applied power electronics conference and exposition (APEC), pp 1326–1331. https://doi.org/10.1109/APEC.2013.6520471
12. Raveendran V, Krishnan NN, Nair MG (2016) Active power compensation by smart park in DC metro railways. In: IEEE 1st international conference on power electronics, intelligent control and energy systems (ICPEICES), pp 1–5. https://doi.org/10.1109/ICPEICES.2016.7853712

13. Kumar KP, Ilango K (2014) Design of series active filter for power quality improvement. In: International conference on electronics, communication and computational engineering (ICECCE), pp 78–82. https://doi.org/10.1109/ICECCE.2014.7086639
14. Sinha R, Vidya HA, Jayachitra G (2020) Design and simulation of hybrid power filter for the mitigation of harmonics in three phase four wire system supplied with unbalanced non-linear load. Int J Adv Sci Technol 29(4):8850–8860
15. Sindhu MR, Nair MG, Nambiar T (2015) Three phase auto-tuned shunt hybrid filter for harmonic and reactive power compensation. Procedia Technol 21:482–489
16. Chen M, Li Q, Roberts C, Hillmansen S, Tricoli P, Zhao N et al (2016) Modelling and performance analysis of advanced combined co-phase traction power supply system in electrified railway. IET Gener Transm Distrib 10(4):906–916
17. Chen M, Liu R, Xie S, Zhang X, Zhou Y (2018) Modeling and simulation of novel railway power supply system based on power conversion technology. In: Proceedings of the 2018 international power electronics conference (IPEC-Niigata 2018-ECCE Asia), Niigata, Japan, pp 2547–2551
18. Ma F, Xu Q, He Z, Tu C, Shuai Z, Lou A, Li Y (2016) A railway traction power conditioner using modular multilevel converter and its control strategy for high-speed railway system. IEEE Trans Transp Electr 2:96–109
19. He X, Peng J, Han P, Liu Z, Gao S, Wang P (2019) A novel advanced traction power supply system based on modular multilevel converter. IEEE Access 7:165018–165028. https://doi.org/10.1109/ACCESS.2019.2949099
20. Teja CH, Ilango K (2014) Power quality improvement of three phase four wire system using modified UPQC topology. Adv Res Electr Electron Eng 1(5):15–21. Print ISSN: 2349-5804; Online ISSN: 2349-5812
21. Sharma A, Sharma SK, Singh B (2018) Unified power quality conditioner analysis design and control. In: IEEE industry applications society annual meeting (IAS), pp 1–8.https://doi.org/10.1109/IAS.2018.8544566
22. Sindhu S, Sindhu MR, Nambiar TNP (2019) Design and implementation of instantaneous power estimation algorithm for unified power conditioner. J Power Electron 19(3):815–826. 1598-2092(pISSN)/2093-4718(eISSN)
23. https://rdso.indianrailways.gov.in/
24. Singh B, Chandra A, Al-Haddad K (2015) Power quality problems and mitigation techniques. Wiley
25. Aseem K, Vanitha V, Selva Kumar S (2020) Performance analysis of DSTATCOM for three phase three wire distribution system. Int Energy J 20(2A):271–278

Intrusion Detection System Using Deep Convolutional Neural Network and Twilio

K. Akhil Joseph Xavier and Gopal Krishna Shyam

Abstract Over the years, human-animal conflict has been a major worry in forest areas and forest zone agriculture fields. As a result, a great number of crops are destroyed, and human lives are put in jeopardy. Farmers and forest officers are working hard to find a solution to the problem. Farmers lose crops, farm equipment, and in some cases, they even lose their lives as well, because of this. As a result, these areas must be regularly maintained to ensure that wild animals do not access agricultural land or human structures. Similarly, large quantities of crops are eaten or destroyed by birds as well. To address the above-mentioned issue, we developed a system that will continuously monitor the field. The proposed system uses a camera system that captures the field's surroundings and is fed as input to the system. Then, it checks if there is an unauthorized intrusion, by analyzing each frame of the input video. The intruder is recognized using a model built using a self-learned deep convolutional neural network (DCNN). In the proposed system, accuracy parameter is used to determine how well the model detects and classifies the animals or bird in real-time video. The proposed system is extremely beneficial to society, particularly farmers and wildlife departments. Many systems exist today, but most of them cause bodily or mental harm to animals. Our technique focuses on preventing wild animals/birds from accessing the fields by monitoring them and not injuring them in any way.

Keywords DCNN · Yolo · Twilio · Darknet

K. Akhil Joseph Xavier (✉) · G. K. Shyam
Department of Computer Science, Presidency University, Bangalore, India
e-mail: kakhiljx@gmail.com

G. K. Shyam
e-mail: gopalkirshna.shyam@presidencyuniversity.in

1   Introduction

Because of the increasing growth of humanity, forest land has been turned into human settlements. As a result, food and water are scarce for wild animals. Deforestation, on the other hand, has forced wildlife to migrate into human areas, putting them at great risk. Both property and people's lives are severely harmed. When we examine the statistics of human-animal conflict over the ages, it is evident that there has been a significant increase. Similarly, because of the radiation emitted by numerous mobile towers, birds are migrating from cities to forest areas which is a consequence of the rapid development in telecom revolutions. We can thereby say that the birds' consumption of crops from the fields increased due to this fact. Typically, farmers employ an electrical fence to protect their fields from animals. These inflict electrocution and cramping on the animals, causing them to act strangely. Animal and human safety are both equally vital. To solve this problem, an intelligent monitoring system is needed that can automatically monitor and recognize the image of an animal/bird entering the field and send an alert message to the user such that the intruder does not get affected either mentally or physically.

The numbers below are taken from a Times of India Article written by Vijay Pinjarkar on September 28, 2019. From 2014 through 2019, the report offers information on human deaths, animal deaths, and crop loss in India. We may deduce from the above-mentioned data that, except for 2019 (because of the covid outbreak), there is a rising trend, with people fighting for their lives. According to data, a system that reduces human-animal conflict and so provides relief to these individuals, especially farmers, is required. This motivated us to the development of the proposed system (Table 1).

Table 1   Animal destruction statistical report

Year	Human kills	Cattle kills	Crop damage
2019	15	2506	8849
2018	36	8311	37,971
2017	50	6909	37,173
2016	53	5961	23,959
2015	46	7812	41,737
2014	42	4496	22,568

2 Proposed System

2.1 Objective

The objective of the proposed system is as follows: (1) To capture the images of the Animals/bird invading the field using a camera. (2) To identify the invaded animal or identify if it is a bird. (3) To play sounds to deviate the animal/bird from the territory. (4) Alert the farmers/forest officers so that they can take suitable action based on the type of intruder.

2.2 Overview

The proposed system consists of a camera trap system (like CCTV, webcam, etc., or even high-resolution cameras) through which the real-time surrounding information is captured continuously and is given as input to the system. We all know that videos are a collection of images (or frames in camera capture perspective). The system reads the video frame by frame with the help of the OpenCV library (commonly known as Open-Source Computer Vision, which is an open-source library for image processing and computer vision applications such as face detection, object tracking, landmark detection, and so on.). Once the frame is read by the OpenCV function, it is examined to find whether there is an intruder present in the frame. The detection of the intruder in the frame is done with the help of a pre-defined 5-layer Deep convolutional neural network (DCNN) model made in combination with the YoloV3 algorithm (You Only Look Once Version 3, which is a real-time object detection system that recognizes specific objects in videos, live feeds, and photos.). If the model detects the presence of an animal, it plays the sound of a wild animal (we have used the tiger sound), whereas if the model detects the presence of a bird, it plays the sound of some dangerous bird (we used the sound of a hawk) to deviate the intruder from the premises. At the same time, the system also alerts the user about the intruder detected so that they can take the necessary action for protecting their premises from them. The alert to the user is done through SMS services. Finally, a log file of the detection details is maintained in the form of a CSV file so that in the future they can have a statistical idea about intruder detection.

2.3 Tools and Methodologies Used

A *Dataset*: In this study, we are not using any pre-available dataset, rather we collect images of targets, pre-process them, and then finally train them. As this project is based on animals/bird detection, the data collected will be the images of various wild animals and some general birds whose presence is being detected

in the field. The source of collecting this data for the project is google. To train the model, we need many pictures for making the model understand the bird and what and how each animal is. In this project, various applications like bulk image downloader and WinZip are used for downloading a larger number of data (images) and then converting them into a zip file, which will be used later in the program.

B *Algorithm*: The algorithm used for intruder detection is a 5-layer CNN algorithm where the five layers are: Convolution, relu, pooling, padding, and fully connected. Yolo V3 in combination with the Darknet framework is also applied in the proposed system. You Only Look Once version 3 (Yolo V3) is a multi-object detection system that uses a convolutional neural network to recognize and classify objects while Darknet is a real-time object detection system built on top of an open-source neural network framework written in C and CUDA.

C. *Tools used*: To carry out the pre-processing of the training data, the Labellmg application was used. The working environment used is Python IDLE (Python 3.6.0). For the intimation of the user, we used the Twilio messenger service provided by using the Twilio Python Helper Library (a library in python that makes us interact with the Twilio API from our Python application). The sound to deviate the animal is played with the help of a function provided by the playsound library and finally, a log file of the detection of animals for performing a statistical analysis is maintained as a CSV file with the help of CSV library where the details like name of intruder detected timestamp, and accuracy is written to it.

2.4 Working of the Proposed System

Like all other systems used in this domain, the proposed system provided by us also has two phases. One is the training phase where the images are trained to get a model and the other is the testing phase where the intruder is detected based on the trained model.

As shown in Fig. 1, the steps involved in the *Training part* are as follows:

A *Data collection*: It refers to the process of collecting the data required for the training of the project and thereby building the model. As this project is based on animals/bird detection, the data collected will be the images of various wild animals and some general birds whose presence is being noticed and is from google.

B *Data Pre-Processing*: The aim of pre-processing is an improvement of the image data that suppresses the unwanted distortions or enhances some image features.
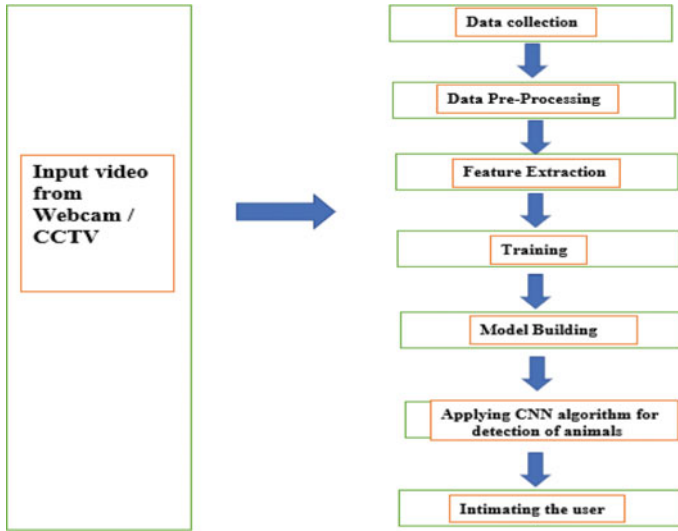
Fig. 1 Major phases of the project

The size and complexity of the image will be reduced during this process. There are basically two processes involved in pre-processing: (1) *Mean Subtraction* and (2) *Scaling*.

Mean Subtraction is done to accomplish two tasks

1. Used to help combat illumination changes in the input images in our dataset.
2. Used to make sure all dataset images are in the same size before feeding to the algorithm.

   Data Scaling or normalization is a process of making model data in a standard format so that the training is improved, accurate, and faster.

C *Feature Extraction*: Feature extraction is the next stage which as the name suggests, is an aspect of an image analysis that focuses on identifying inherent features of the objects present within an image. These inherited features are then converted into pixel values later during training and then these pixel values are compared with the image pixel value (input) and based on this comparison classification is performed.

D *Training and Model Building*: Once the feature extraction is done, we move forward to the training and model building part. In the project we make use of the CNN algorithm and as we all know CNN algorithm works fine only with image files. Since our input is live video, we cannot apply CNN algorithms directly. So, we make use of the YoloV3 method in combination with Darknet to build the model before feeding it to the CNN algorithm.

The steps involved in the *Testing part* are as follows

A *Image Acquisition*: It is the initial step involved in the testing part through which the real-world sample is recorded in its digital form by using a digital camera. The input video is first converted into frames before giving to the CNN algorithm. This is done by reading the video frame by frame (by looping) using read fn in OpenCV. The image is captured, scanned, and converted into a manageable entity.

B *Image Pre-processing and Feature Extraction*: OpenCV provides a function (cv2.dnn.blobFromImage) to facilitate image pre-processing and feature extraction for deep learning classification.

The cv2.dnn.blobFromImages function returns a blob, i.e., input image after mean subtraction, normalizing, and channel swapping. Blob stands for Binary Large Object and refers to the connected pixel in the binary image.

C *Classification*: It is in this stage that the detected intruder is identified and classified as birds or animals like a tiger, lion, etc. The blob created after pre-processing and feature extraction will be fed into the algorithm and the algorithm compares this blob with the model and based on that it predicts the detection of the intruder., i.e., based on the similarities in the pixel value obtained after pre-processing and feature extraction with the respective pre-defined pixel value of the corresponding animal/bird, it classifies the intruder as birds/animal.

D *Intimation, playing sound, and Log file creation*: This is the final step in the testing part where the user gets the alert about the intruder. The intimation to the authorities is done through Twilio messenger where the alert will be sent in the form of SMS with the aid of the internet. Similarly, when an intruder is detected, the system plays a sound to keep the intruder away. This can be connected to a speaker or an amplifier to enhance the sound frequency if required. Finally, a CSV file is created to store all the log/history of the intruder detected with names, accuracies, and timestamps as parameters so that we can fetch the details of the intruder detected within a timeline. In this study, we have provided the user to choose from three timeline analyses namely date wise, month wise, and year-wise statistics. Based on the user selection, the user will get the details of the value counts of the intruder along with their names, accuracy, and timestamp in the given timeline so that the user gets to know about the seasonal arrival of the intruder and the frequency of their arrivals.

2.5 System Implementation

The system implementation for the proposed system is shown as a flowchart in Fig. 2:
The steps involved in the implementation are:

1. Collecting images of animals as a database which is used for training of our program
2. Based on the database, a model is being created for testing.
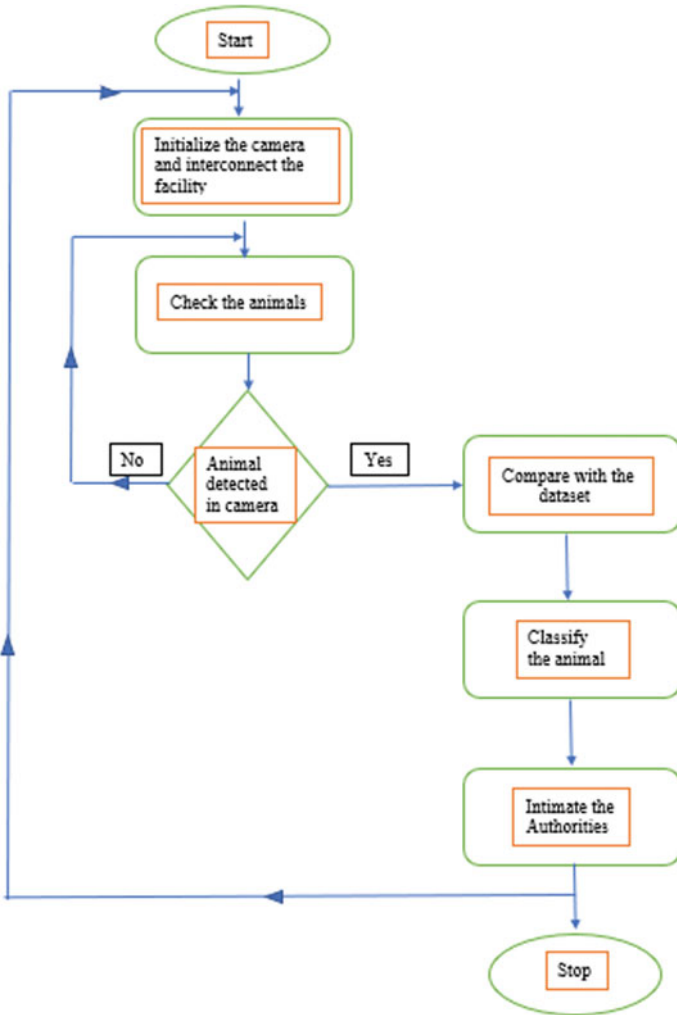
Fig. 2 System implementation for the proposed system

3. The Image/video acquisition from the camera is done and is fed as input to the system.
4. Once input is fed to the system, the conversion of video to frames is carried out.
5. We use the imread function to read the camera captured frame (image) if it is fetching video stored in a hard disk or cv2.videocapture() if the camera is directly connected to the computer. After this Preprocessing is done on that image/frame.
6. After pre-processing of the camera captured image, a comparison of this frame with the model created is carried out., i.e., the system checks if matching is found or not.

7. If matching is found, intimation to the concerned person about the animal is done. Else it continues from step 3.

3 Experiment Results and Discussions

In the output of the experiment, the animal along with its accuracy of prediction and time stamp is displayed. When the animal was detected, a tiger sound was played, and the details were written to CSV. The statistical analysis was done successfully on the log file (CSV file) where monthly, yearly, and date-wise statistical analysis could be done. The intimation to the user was successfully received in the form of an SMS.

The experimental results of the system are depicted in Figs. 3, 4, 5 and 6 as follows:

Output 1

See Fig. 3.

Output 2

See Fig. 4.

Output 3

See Fig. 5.

Output 4

See Fig. 6.
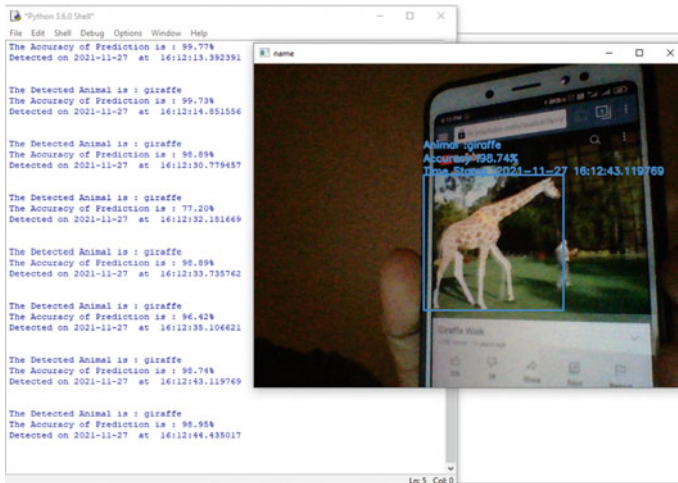
Fig. 3 Detecting animal (giraffe) from the live video
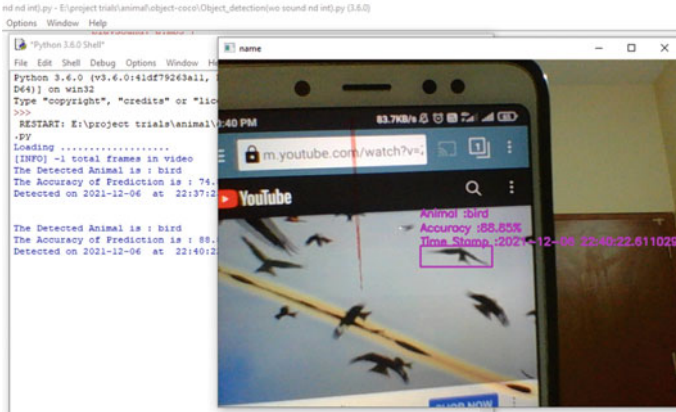
Fig. 4 Detecting bird from live video
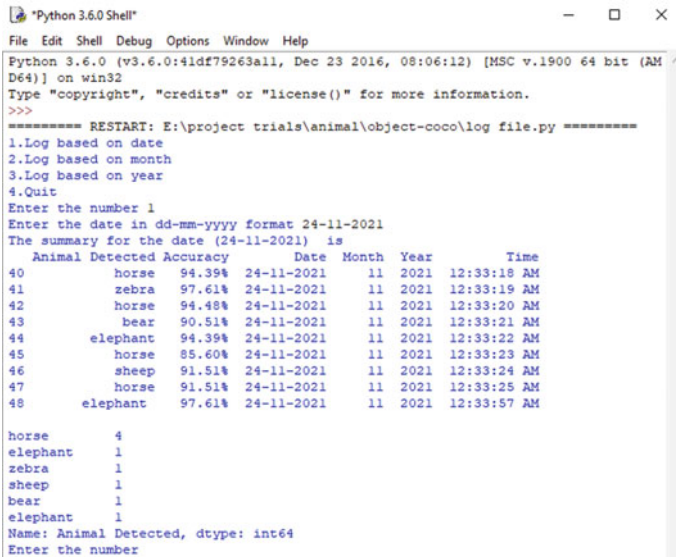
Fig. 5 Log file experimental results

4 Conclusion and Future Scope

In this study, the model was trained for nine animals and some general birds. Three animals and birds were tested in the system. The three animals are giraffes, sheep, and horses. The model predicted these animals accurately with an accuracy of almost 90%. Similarly, the model also predicted the birds accurately with an accuracy of almost 75%. The accuracy of birds was low when compared to the accuracies of

Fig. 6 Intimation to the user via SMS
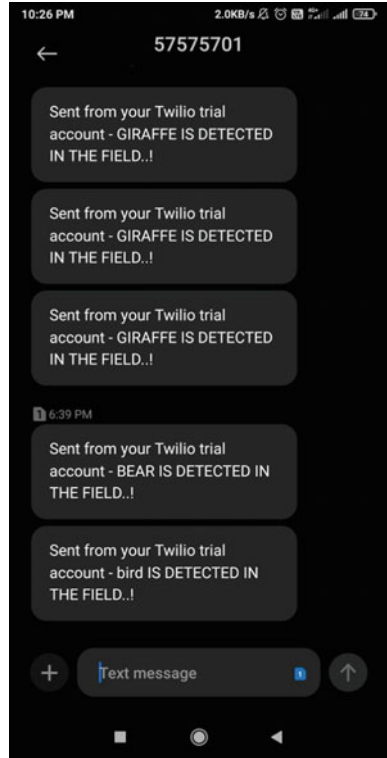
animals as birds fly faster and hence their images in each frame will have fewer details. However, if we use high-resolution cameras, the accuracy could be improved. As soon as the system detected the presence of the animal, the sound of the tiger was played. Similarly, the sound of a hawk was played in the detection of birds. The details of detection like intruder name, accuracy, and timestamp were successfully written to a CSV file. Based on the CSV file created, the log of the data was maintained. We fetched the details of intruder detected within a timeline along with their accuracies and was shown in the results. By maintaining the log file, the user can be aware of the frequency of intruder arrival and seasonal arrival of the intruder if any.

Currently, we have taken into consideration of only nine animals and some general birds to the training set as it takes relatively a very long time, and mandatory GPU is required to train these intruders. We used google colab to train the model weights, which provides 12 h of GPU support. By using google colab pro, the GPU support could be extended up to 24 h so that we can extend the model to specific birds as well. In doing so, the farmers can get to know specifically which bird is intruding and adapt individual strategies to fly them away. Similarly, the system could be modified with hardware support as well by using an Arduino or raspberry pi along with some cameras attached to it to fetch the input video. The proposed system could also be further extended to moving videos as well such that the hardware will be connected

to a drone to fetch the live videos. In doing so, the system can continuously monitor all the territories in the field rather than concentrating on a single area.

Bibliography

1. Singh A, Pietrasik M, Natha G, Ghouaiel N, Brizel K, Ray N (2020) Animal detection in man-made environments. In: Proceedings of IEEE conference on 2020 IEEE winter conference on applications of computer vision (WACV)
2. Schindler F, Steinhage V (2021) Identification of animals and recognition of their actions in wildlife videos using deep learning techniques. Proc Sci Dir J Ecol Inform 61
3. Shetty H, Singh H, Shaikh F (2021) Animal detection using deep learning. Proc Int J Eng Sci Comput (IJESC) 11(06)
4. Kruthi HI, Nisarga AC, Supriya DK, Sanjay CR, Mohan Kumar KS (2021) Animal detection using deep learning algorithm. Proc Int J Adv Res Comput Commun Eng (IJARCCE) 10(6)
5. Kiruthika K, Janani Sri AR, Indhumathy P, Bavithra V (2020) Automating the identification of forest animals and alerting in case of animal encroachment. Proc Int Res J Eng Technol (IRJET) 07(03)
6. Banupriya N, Harikumar S, Saranya S, Palanasamy S (2020) Animal detection using deep learning algorithm. Proc J Crit Rev 7(1)
7. Udaya Shalika AWD, Seneviratne L (2016) Animal classification system based on image processing and support vector machine. Proc J Comput Commun 4
8. Clune J, Packer C, Nguyen A, Kosmala M, Swanson A, Palmer MS, Norouzzadeh MS (2018) Automatically identifying, counting, and describing wild animals in camera trap images with deep learning. Proc Nat Acad Sci USA
9. Prajna P, Soujanya B, Divya S (2018) IOT-based wild animal intrusion detection system. Proc Int J Eng Res Technol (IJERT). ISSN: 2278-0181, ICRTT—2018 Conference)
10. Andavarapu N, Vatsavayi VK (2017) Wild animal recognition in agriculture farms using W-COHOG for agro-security. Proc Int J Comput Intell Res 13(9). (ISSN 0973-1873)

A Neural Network-Based Cardiovascular Disease Detection Using ECG Signals

C. Venkatesh, M. Lavanya, P. Naga Swetha, M. Naganjaneyulu, and K. Mohan Kumar Reddy

Abstract Heart disease is also known as cardiovascular disease (CVD). It is a type of illness that affects the heart and blood vessels (veins and arteries). Cardiovascular disease (CVD) is the most major cause of mortality in both men and women around the world. The enormous increase of these diseases and associated consequences, as well as their high expenses, have communities will be harmed, and the international community will be burdened financially and physically. Cardiovascular disease cannot be easily anticipated by medical practitioners since CVD is a challenging process that needs experience and advanced understanding. Though real-life consultants can forecast disease with more tests and it takes long time to process, their prognosis may be wrong due to a lack of expert knowledge. As a result, employing effective preventative techniques is critical. Heart sound signals provide crucial information about a person's health. Previous research has revealed that single-channel cardiac sound waves are utilized to identify cardiovascular disorders. The precision of a single-channel cardiac sound signal, on the other hand, is insufficient. A novel method for fusing multi-domain features of multi-channel cardiac sound signals and convolutional neural network is described in this paper.

Keywords Cardiovascular disease · ECG signals · Clustering · Convolutional neural network

C. Venkatesh (✉) · M. Lavanya · P. Naga Swetha · M. Naganjaneyulu · K. Mohan Kumar Reddy
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: cvs@aitsrajampet.ac.in; venky.cc@gmail.com

1 Introduction

Coronary artery disease (CAD) is the leading cause of death in people with heart disease, and the problem is getting worse [1]. As a result, simple and effective methods for detecting CAD in large populations are urgently needed. Coronary angiography (CAG) is used for diagnosing coronary artery disease (CAD) [2]. However, because of its invasiveness and expensive cost, it is not suited as a regular inspection procedure for early screening. According to medical studies, blood travelling through a stenosis in a blood vessel impacts the channel's wall, causing turbulence. Murmurs in cardiac sound signals can be caused by turbulent air [3]. Because it is a semi-approach of detection, analysis of heart sound has the possibility to become a cost-effective evaluation tool for detecting CAD early [4].

CVD is a type of illness that impacts the heart and the blood vessels (veins and arteries). In the early twentieth century, heart diseases were responsible for 10% of all deaths [5], and by the late twentieth century, the death rate from cardiac diseases had increased by 25%. According to the World Health Organization, cardiovascular disease remains the major cause of death worldwide. In 2004,17.1 million people died from cardiovascular disease, which was 29% of all fatalities globally. CAD is the most common disease in CVDs, accounting for 7.2 million deaths. Cardiac trauma was responsible for 5.7 million fatalities. Low and moderate-income nations are disproportionately impacted; 82% of deaths due to CVD occur in these countries, and the percentage is comparable in both men and females. It is estimated that over 23,600,000 people will die between now and 2030, and the majority of them will die from heart disease and stroke. Because of changes in lifestyle, dietary habits, and working environment, the eastern Mediterranean region will see the most fatalities, with Southeast Asia accounting for the majority of deaths. As a result, according to WHO data, it is critical to use precise methodologies and efficient periodic cardiac examinations to identify heart disorders [6].

Heart disease is one of the major causes of mortality. It is tough for medical practitioners to forecast since it is a complex undertaking that needs competence and a greater level of information. Cardiovascular disease is the leading cause of death in humans all over the world. Although real-life consultants can forecast disease using a large number of tests and a long processing period, their predictions are occasionally wrong due to a lack of expert knowledge [7]. In the realm of CVD prevention, the most pressing issue is determining whether or not a person is likely to develop the condition. Many works are motivated by rising annual CVD patient mortality and the accessibility of vast amounts of patient records to retrieve useful information [8].

2 Literature Review

Nourmohammadi-Khiarak et al. [9] suggested a new technique for heart disease detection that uses an optimization algorithm in feature selection, data mining, and K-nearest neighbour, with an accuracy of 88.25%.

Abdeldjouad et al. [10] postulated a method for predicting and diagnosing heart disease. They have performed data mining techniques utilizing artificial intelligence and machine learning techniques, assisting in medical data analysis with an accuracy of 80%.

Liu et al. [11] proposed a new method for detecting coronary artery disease. The accuracy achieved after adding entropy and cross entropy features to multi-domain feature fusion of multi-channel heart sound signals with SVM for classification is 90%.

Akhil Jabbar et al. [12] proposed a method for classifying heart disease based on k-nearest neighbour and the genetic algorithm. They used KNN in conjunction with a general algorithm for successful classification and data mining to evaluate medical data and extract useful information with an accuracy of 88%.

Pathak et al. [13] proposed a method to detect coronary artery disease in a noisy environment using phonocardiogram signals, and they used the imaginary part of cross power spectral density to acquire the spectral range of heart sounds because it is non-responsive to zero time-lag signals, and spectral features obtained from ICPSD are classified in a machine learning framework, and this method obtains 74.98% accuracy.

The accuracy attained by all of these approaches is less than or equal to 90%. As a result, we present a novel approach for detecting cardiovascular disease using optimized multi-channel cardiac sound signals to increase accuracy.

3 Methodology

As depicted in Fig. 1, the proposed system initiates its operation by acquiring the ECG samples of heart from different nodes and reshapes for proper sequence by properly buffering the ECG Pulses. The buffered ECG pulses are suitably fused by multimodal fusion frame work to combine all the pulses of similar magnitude and modalities. Then, the fused ECG pulses are clustered into different classes based on their pulse magnitudes by k-means. The classification features are extracted from the fused and clustered ECG pulses and are subjected to the abnormalities detection and classification by the convolutional neural network. The detected abnormalities are again clustered into variant groups for disease classification. Finally, the classification results are conveyed with psychovisual and quantitative parameters such as sensitivity, specificity, accuracy, PSNR, and MSE.

```
┌─────────────────┐      ┌─────────────────┐      ┌─────────────────┐
│ Heart ECG signals│─────▶│ Buffering of ECG│─────▶│ Fusion of buffered│
│  (Multi Channel) │      │    samples      │      │   ECG pulses    │
└─────────────────┘      └─────────────────┘      └─────────────────┘
                                                            │
                                                            ▼
┌─────────────────┐      ┌─────────────────┐      ┌─────────────────┐
│ Construct a class│◀─────│ Extract ECG features│◀──│ Clustering of ECG│
│ oriented feature │      │  from fused &   │      │ pulses with different│
│     vector      │      │ clustered ECG pulses│   │    magnitudes   │
└─────────────────┘      └─────────────────┘      └─────────────────┘
        │
        ▼
┌─────────────────┐      ┌─────────────────┐      ┌─────────────────┐      ┌─────────────────┐
│   Detected      │─────▶│  Cluster the    │─────▶│    Disease      │─────▶│    Output       │
│  Abnormalities  │      │   abnormal      │      │ Classification  │      │  (Attributes)   │
│     (CNN)       │      │    pulses       │      │                 │      │                 │
└─────────────────┘      └─────────────────┘      └─────────────────┘      └─────────────────┘
        ▲
        │
┌─────────────────┐
│ Trained Dataset │
└─────────────────┘
```
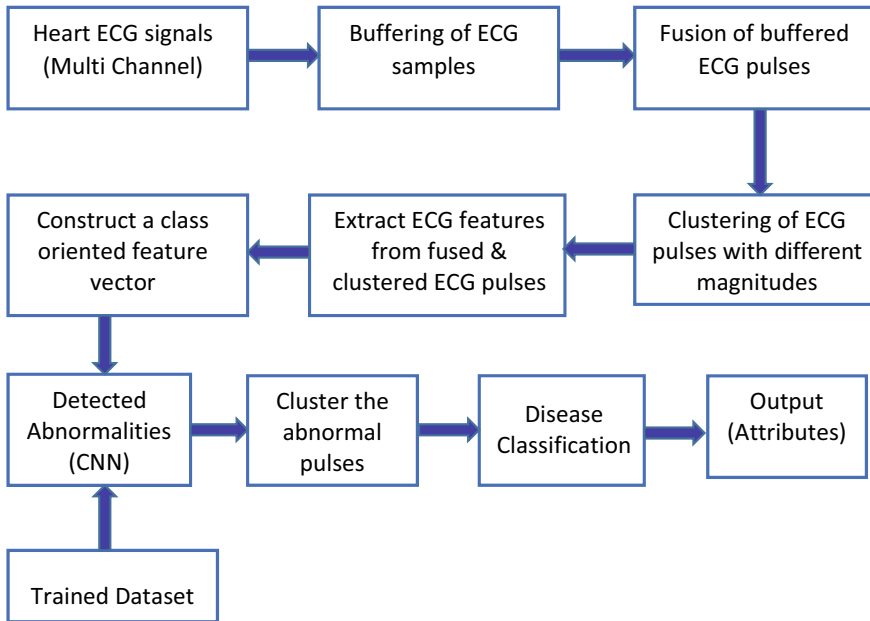
Fig. 1 Proposed block diagram

3.1 ECG Signals

The electrocardiogram (ECG) monitors the electrical signals produced by cardiac myocytes and determines whether the heart is beating normally or abnormally. ECG signals are used to diagnosis the cardiovascular disorders. Normal ECG signals have tiny amplitude ranging from 10 V to 5 mV and vary in time. They have a fairly standard value of 1 mV and frequency range ranging from 0.05 to 100 Hz [14]. Typically, the lead channel obtains it from body surface electrodes. Because of its non-intrusiveness, output visualization, low cost, convenience, and adaptability, the acquisition method is widely used [15–17]. In ECG, the P, QRS, and T waves are present.

P Wave P wave indicates the activation of the right and left atria in a specific order, and it is normal to witness right and left atrial activity with notched or biphasic P waves.

QRS Complex The QRS complex denotes ventricular depolarization and represents the electrical impulse as it travels through the ventricles. The QRS complex, like the P wave, begins just before ventricular contraction.

T Wave The T wave on the ECG represents ventricular myocardium repolarization (T-ECG). Its morphology and duration are frequently used to diagnose pathology and determine the risk of life-threatening ventricular arrhythmias.

3.2 Buffering

Buffering is a storage technique. The ECG samples that have been collected have been pre-processed. All of the signals are reshaped to ensure that they are in the correct order.

3.3 Multimodal Fusion

Fusion is the process of combining two or more objects into a single entity. A multimodal approach can be used to conduct fusion. Multimodal fusion integrates data from several sources and combines it into a single command [18].

Fusion can be done in the following ways:

1. Recognition-based.
2. Decision-based.
3. Multi-level hybrid fusion.

3.4 Clustering

K-means clustering is used. It is difficult to locate and merge comparable data pieces in larger data sets. In this work, k-means clustering is used to locate and merge such data fragments. The goal is to divide n observations into k groups. Each observation is a member of the cluster, with the nearest mean acting as the cluster's prototype. The k-means clustering technique primarily accomplishes two goals. These are the following: Iteratively determines the optimal value for k-centre points or centroids and assigns each data point to the nearest k-centre. A cluster is formed by data points that are close to a specific k-centre. As a result, each cluster contains data points with certain commonality and is isolated from the others. K is a numerical value and it ranges from 1 to N and we select randomly and minimum value is 3 and maximum value we can take any odd number.

3.5 Extraction

Feature extraction is the important step in classification [19]. A type of dimensionality reduction is feature extraction. The image's large number of pixels is efficiently represented. It aids in the reduction of redundant data in the data set. Extract ECG characteristics such as maximum value, minimum value, range, and so on. Clustered ECG pulses are used for feature extraction. These features are fed to convolutional neural network for classification.

3.6   Classification by Convolutional Neural Network

Classification is the systematic arrangement of things on the basis of certain similarities. One type of artificial neural network is convolutional neural networks (CNNs). CNN is used for image recognition and processing. It was created with the sole purpose of processing pixel data [20]. Convolutional neural networks (CNNs) are data processing systems that use multiple layers of arrays to process data. CNN follows the biological principle of structure replication. It is capable of detecting patterns in a variety of settings. CNN is composed of several layers [20]. The input is routed through a number of layers. A pooling layer is a type of down sampling layer that is not linear. Each and every node is connected in fully connected layers. CNN is distinct from other neural networks in that it accepts input in rather than focusing on feature extraction like other neural networks, and it uses a two-dimensional array and operates directly on the images. The working procedure of CNN is as depicted in Fig. 2 and as follows:

Convolution Layer The first layer, known as the convolution layer, is used to extract the features maps from the source images (while retaining pixel relations) (CL). If the CL is nonlinear, the second layer is an activation layer that comes after it. In this layer, the rectified linear unit (ReLU) function is the most versatile and common activation function. ReLU is always represented by a pair of zeros (0) and ones (1).

Pooling Layer The layers that appear after the CNN declaration are known as pooling layers or convolutional neural networks. It converts a feature map generated by convolutional networks into a condensed feature map using the user's input. By combining neurons from previous layers, pooling layers aid in the creation of new layers [21].
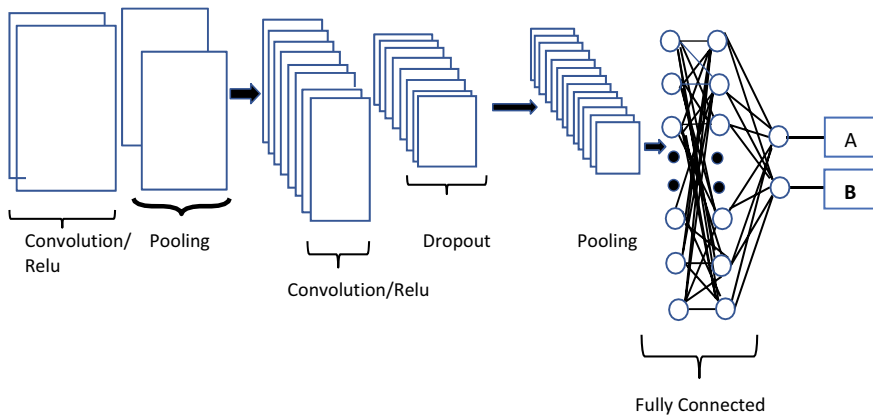
Fig. 2 CNN phase classification

Fully Connected Layer To get a high-level idea of the qualities, this layer primarily functions as an integrator. The fully connected layer receives data from the convolutional and pooling layers. Each neuron in this layer is linked to every neuron in the previous layer, but no neurons in the same layer are linked.

In this work, we use three modules they are input modules, hidden modules, and output modules. Input modules are used to take input, and in the hidden modules, we use 15 layers and accuracy is based on this hidden modules and output modules provide the output. The entire process occurs in the hidden layer.

3.7 Train Data Set

The training data is a collection of information. It is first used to train the algorithm or programme. Better outcomes have been discovered. In this work, 50–100 data sets are used.

4 Results

The proposed system detects cardiovascular by using cardiac sound signals. To detect and classify the cardiovascular diseases, initially the ECG signals are acquired. The ECG signal provides extensive information about the heart's structure and function. ECG signals which were taken from various nodes of the patient are shown in Fig. 3.

To process all 4 signals at a time, it is complicated. So, to reduce the complexity of the process, the 4 ECG signals of patients are fused. A multimodal fusion strategy enhances heartbeat detection by utilizing additional information present in the various physiological signals. The fused ECG signals of patient are shown in Fig. 4.

After selecting the optimal features, then those features are extracted by clustering technique. Clustering is the process of dividing a population or set of data points into multiple groups so that data points in the same group are more similar than data points in other groups. The clustered ECG pulses are as shown in Fig. 5.

Figure 6 shows the training phase of neural network and the performance of the corresponding training phase is as shown in Fig. 7.

The total number of epochs used to train the features which are obtained from clusters is shown in Fig. 8, and the corresponding histogram is as shown in Fig. 9.

The confusion matrix is shown in Fig. 10. The number of correct predictions for each class falls along the diagonal matrix. If the ROC curve was near to one, then it indicates the system is perfect as shown in Fig. 11.The statistical attributes obtained from the proposed system are given below.
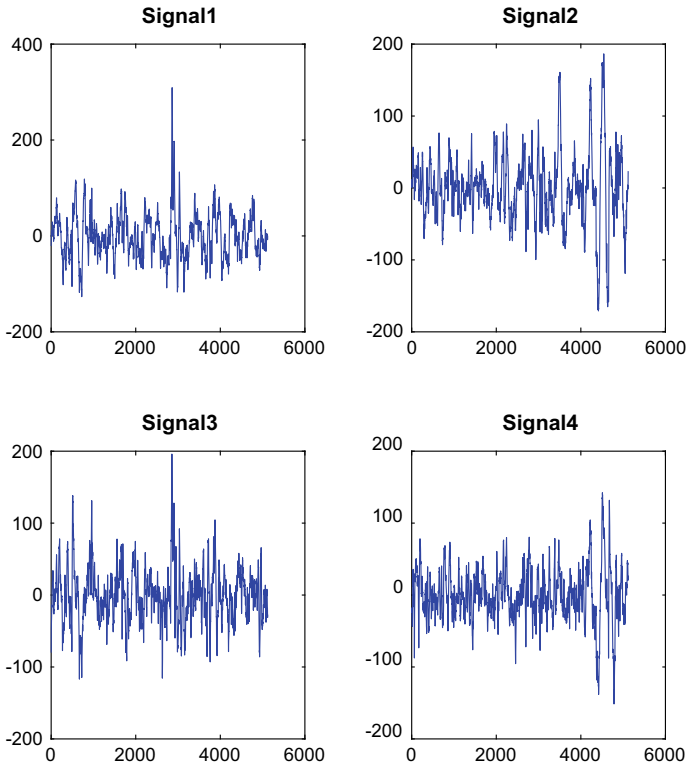
Fig. 3 ECG signals

Detected Heart Disease: HEART VALVE
Accuracy of CVD Detection: 98.000000 Inconclusive Rate of CVD Detection:
0.000000 Classified Rate of CVD Detection: 100.000000 Sensitivity of CVD
Detection: 71.875000
Specificity of CVD Detection: 62.500000
Positive Predictive Value of CVD Detection: 88.461538
Negative Predictive Value of CVD Detection: 35.714286
Positive Likelihood of CVD Detection: 191.666667
Negative Likelihood of CVD Detection: 45.000000
Prevalence of CVD Detection: 80.000000.

5 Conclusion

In this work, convolutional neural network is used for detecting different cardio-vascular diseases by recording electrocardiogram signals and fusion is performed by multimodal fusion for integrating. Feature of multi-domain the integration of
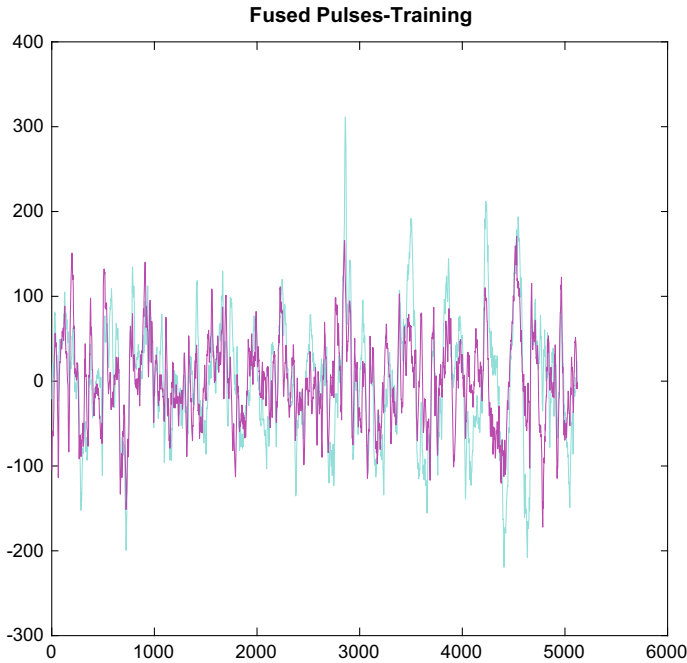
Fig. 4 Fused ECG signals
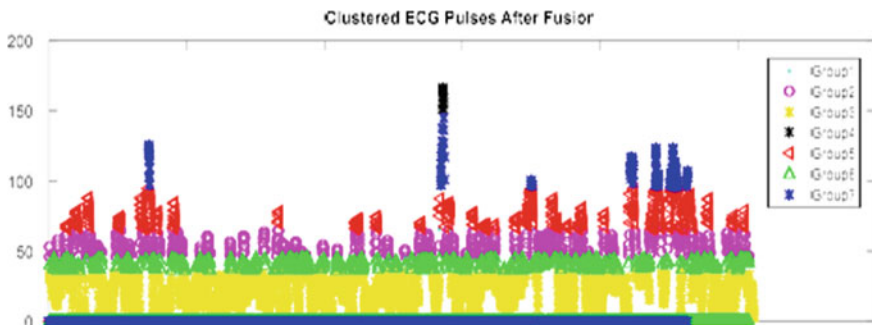
Fig. 5 Clustered ECG pulses

multi-channel heart sound waves can provide extra information that can aid in the prevention and treatment of cardiovascular illnesses. The fusion signals clustered by using k-means clustering technique and feature extraction from the fused and clustered electrocardiogram pulses and class-oriented feature vector created and features are subjected to the abnormalities detection and classification by the convolutional neural network. The detected abnormalities are again clustered into variant groups
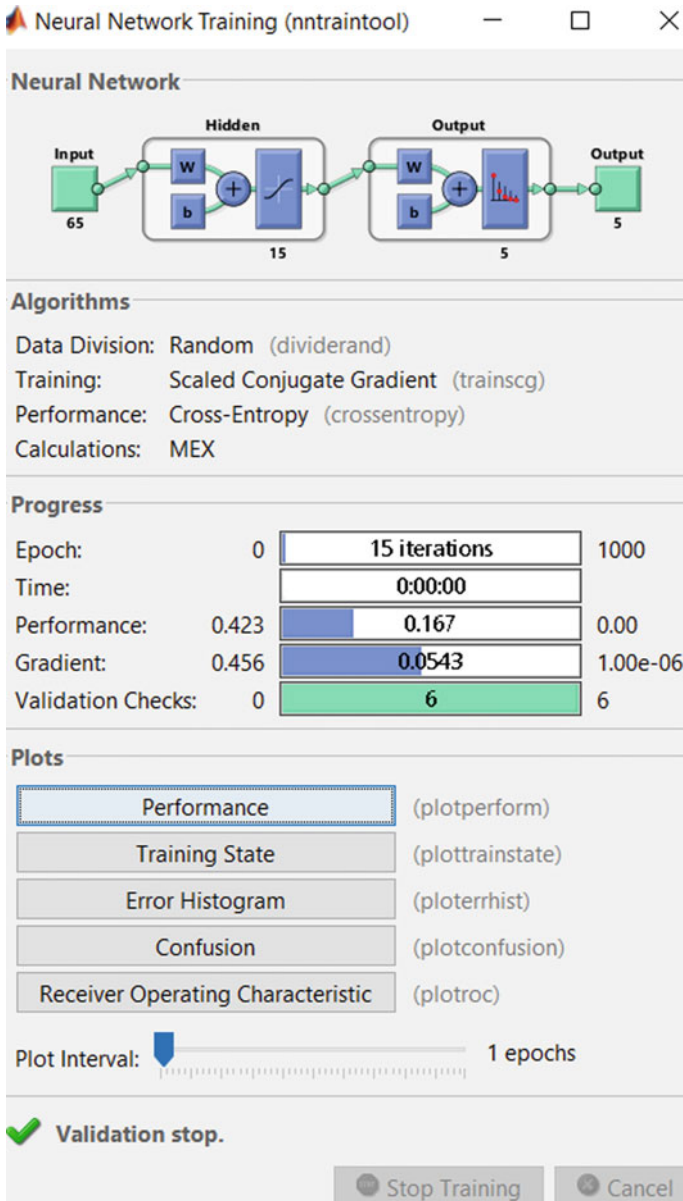
Fig. 6 Neural network
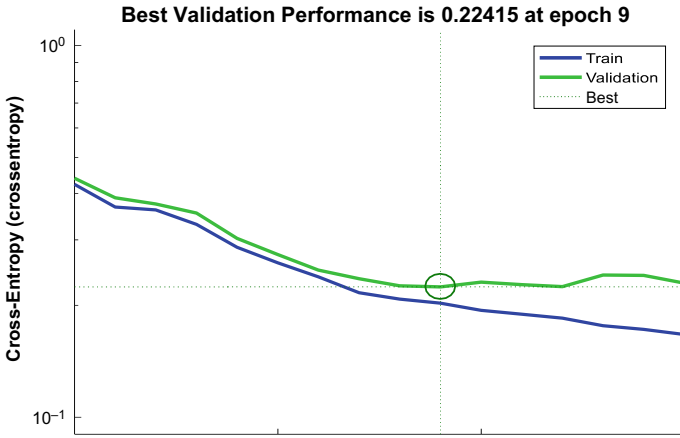
Fig. 7 Performance
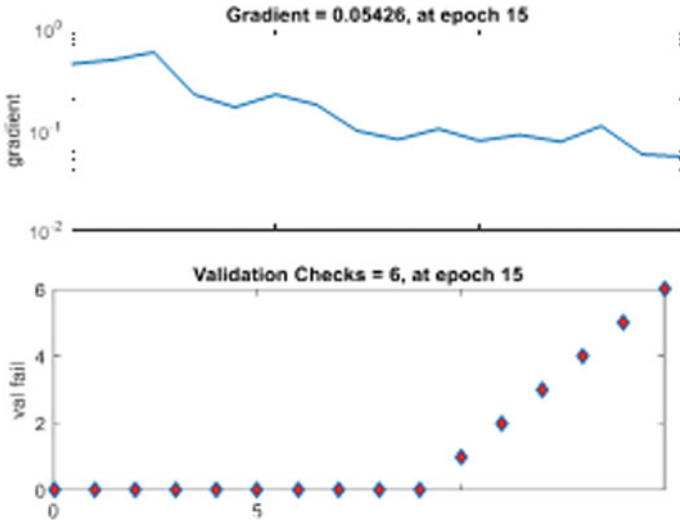
Fig. 8 Training state

for disease classification. The classification results are conveyed with psychovisual and quantitative parameters.

Fig. 9 Histogram
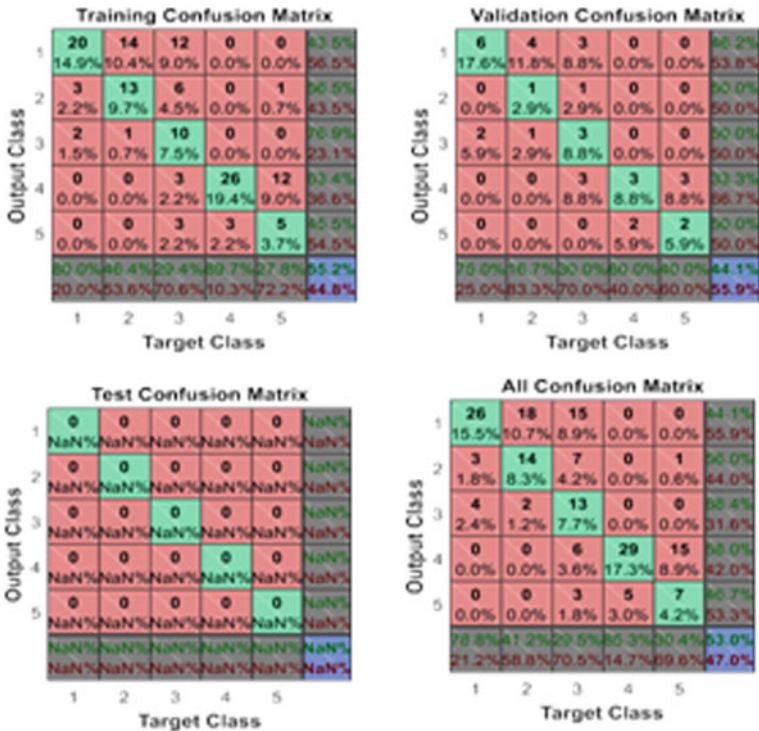
Fig. 10 Confusion matrix
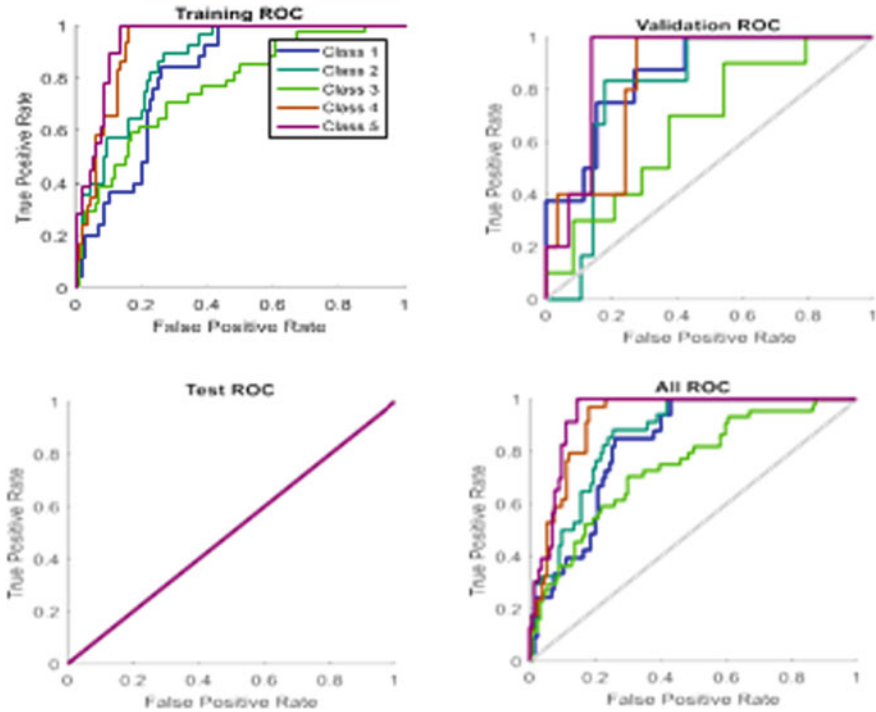
Fig. 11 ROC curve

References

1. Alonso MS, Bittencourt CW, Callaway AP, Carson AM, Chamberlain AR, Chang S, Cheng SR, Das et al (2019) The 2019 update on heart disease and stroke statistics from the American Heart Association. Circulation 139:e56–e528. [CrossRef] [PubMed]
2. Yoshida H, Yokoyama K, Maruvama Y, Yamanoto H, Yoshida S, Hosoya T Coronary artery calcification using coronary angiography to diagnose stenosis in haemodialysis patients (CAG). Nephrol Dial Transplant 1451–1452. [CrossRef]
3. Rahalkar K, Semmlow J (2009) Coronary artery disease can be detected acoustically. Biomed Eng Annu 9(4):449–469. [CrossRef][PubMed]
4. Mahnke C (2009) Automated heartsound analysis/computer-aided auscultation: a cardiologist's perspective and future development suggestions. In: Proceedings of the IEEE engineering in medicine and biology society's 2009 annual international conference, Minneapolis, Minnesota, pp 3115–3118
5. Deekshatulu B, Chandra P (2013) Classification of heart disease using K-nearest neighbor and genetic algorithm. Procedia Technol 10:85–94
6. World Health Organization (2000) The world health report 2000: health systems: improving performance. World Health Organization
7. Pouriyeh S, Vahid S, Sannino G, Pietro GD, Arabnia H, Gutierrez J (2017) A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease. In: IEEE symposium on computers and communication, Heraklion, Greece, pp 1–4

8. Safdari R, Samad-Soltani T, GhaziSaeedi M, Zolnoori M (2018) Evaluation of classification algorithms versus knowledge-based methods for differential diagnosis of asthma in iranian patients. Int J Inform Syst Serv Sect 10(2):22–26

9. Nourmohammadi-Khiarak J, Feizi-Derakhshi M-R, Berouzi K, Mazaheri S, Zamani-Harghalani Y, Tayebi RM (2019) New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection. Springer

10. Abdeldjouad FZ, Brahami M, Matta N (2020) A hybrid approach for heart disease diagnosis and prediction using machine learning techniques. Springer

11. Liu T, Li P, Liu Y, Zhang H, Li Y, Jiao Y, Liu C, Karmakar C, Liang X, Ren X, Wang X (2021) Detection of coronary artery disease using multi-domain feature fusion of multi-channel heart sound signals. MDPI Entropy

12. Akhil Jabbar V, Deekshatulu B, Chandra P (2013) Classification of heart disease using Knearest neighbor and genetic algorithm. Procedia Technol 10:85–94

13. Pathak A, Samanta P, Mandana K, Saha G (2020) An improved method to detect coronary artery disease using phonocardiogram signals in noisy environment. Appl Acoust Sci Dir

14. Singh BN, Tiwari AK (2006) Optimal selection of wavelet basis function applied to ECG signal denoising. Digit Signal Process 16:275–287. [CrossRef]

15. Baloglu UB, Talo M, Yildirim O, Tan RS, Acharya UR (2019) Classification of myocardial infarction using multi-lead ECG signals and deep CNN. Pattern Recogn Lett 122:23–30. [Source: Google Scholar] [CrossRef]

16. Black N, D'Souza A, Wang Y, Piggins H, Dobrzynski H, Morris G, Boyett MR Circadian rhythm of cardiac electrophysiology, arrhythmogenesis, and underlying mechanisms. Heart Rhythm 16(3):298–307. [Source: Google Scholar] [CrossRef]

17. Dietrichs ES, McGlynn K, Allan A, Connolly A, Bishop M, Burton F, Kettlewell S, Myles R, Tveita T, Smith GL (2020) Cardiovascular research. 116:2081–2090. [Source: Google Scholar] [CrossRef] [PubMed]

18. Zhu H, Wang Z, Shi Y, Hua Y, Xu G, Deng L (2020) Multimodal fusion method based on self-attention mechanism, Hindawi

19. Kutlu Y, Kuntalp D (2012) Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients. Comput Methods Programs Biomed 105(3):257–267

20. Jiang A, Jafarpour B (2021) Inverting subsurface flow data for geologic scenarios selection with convolutional neural networks. Adv Water Resour 149:1–17

21. Xu H, Wang D, Deng Z et al (2020) Application of remote sensing fuzzy assessment method in groundwater potential in Wailingding Island. J Supercomput 76(5):3646–3658

Detection and Classification of Lung Cancer Using Optimized Two-Channel CNN Technique

C. Venkatesh, N. Sai Prasanna, Y. Sudeepa, and P. Sushma

Abstract Lung cancer is one among the most dangerous diseases that cause deaths across the globe. Within the respiratory organs, cells can display through the bloodstream or humour fluid that surrounds through the lung tissues. Respiratory organ affected 297 (13.1%) males and fifty-nine (2.5%) females. With a male-to-female magnitude relation of 5:1, this respiratory organ abnormality stands second among males and tenth among females. Here we go for unique hybrid model planned for detection and classification. The first stage starts with taking a group of CT pictures from the public and private. CT pictures are pre-processed from the info base, and it removes the noise through median filter. Later the image is segmented by RCNN technique in conjunction with the improvement of cuckoo search, and also, the feature extraction is performed and classified by using the RCCN with the two-channel CNN technique. Wherever the dimensions and placement of unwellness are detected. Finally, the detected unwellness image classifies it's benign or malignant. Later we obtain parameters like accuracy, sensitivity, specificity, entropy and mean square error rate (MSE).

Keywords Optimization · RCNN · Segmentation · CNN · Lung cancer

1   Introduction

Cancer is one of the most dangerous diseases that can cause death. A quarter of the population will be diagnosed with cancer at some time in their lives, with an annual mortality rate of 171.2 per 100,000 people [1]. When it comes to carcinoma, the death rate is exceedingly high, and even when the cancer is discovered, the survival percentage is extremely low. Carcinoma is the most common and deadliest cancer. The death rate of cancer patients could be decreased with early identification and analysis. The earlier a person is detected, the higher chances of surviving. There is a

C. Venkatesh (✉) · N. Sai Prasanna · Y. Sudeepa · P. Sushma
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: cvs@aitsrajampet.ac.in; venky.cc@gmail.com

progressive increase in the number of cancer deaths once a year. Cigarette smoking causes lung cancer in 85% of men and 75% of women, according to research [1].

Lung cancer symptoms include coughing that gets worse, chest pain, weight loss, shortness of breath, bloody coughing and weariness [2]. Hospitals are employed to obtain CT scan images and blood samples. Computerized tomography results are less noisy than MRI and X-ray readings [3]. As a result, detecting lung cancer early and precisely is crucial. Many studies have used machine learning models to detect and categorize lung computed tomography (CT) pictures [4, 5].

Various computer-assisted detection (CAD) methods, such as convolutional neural networks (CNNs), have proven classification performance on medical images in recent years. CT scans are frequently used in screening strategies to reduce high mortality and gather information about lung nodules and their environs. In the medical field, CNNs have demonstrated remarkable performance in applications such as vision and graphics in CAD systems to accurately diagnose lung cancer from CT scans [6].

A CNN model with three convolutional layers, three pooling layers, and two fully connected layers was employed for classification [7, 8]. Two-dimensional (2D) CNN has been utilized with promising results in image classification, facial recognition and natural language processing, to name a few disciplines [9]. The current CAD design calls for the training of a large number of parameters, but parameter setup is challenging, and thus, the parameters must be fine-tuned to increase classification accuracy [10].

CNN was used to classify the object's presence within that region. The problem with this method is that the objects of interest may be in different spatial locations and have varying aspect ratios within the image. As a result, you'd have to pick a big number of regions to compute, which may take a long time. As a result, algorithms such as RCNN, YOLO and others have been created to discover these events quickly.

The original purpose of RCNN was to take an input image and output a set of bounding boxes, each of which contained an item as well as the object's category (e.g. car or pedestrian). RCNN has recently been extended to perform other computer vision tasks. RCNN starts by extracting regions of interest (ROI) from an input image using a process called selective search, where each ROI is a rectangle that may indicate the boundary of an item in the image [11].

The input image is collected from the test dataset, and it undergoes pre-processing to resize and transform it. After that, we'll use a hybrid technique of segmenting the tumour image with RCNN and cuckoo optimization to find the portion component. The type of tumour is detected, and tumour characteristics are determined, based on the segmented output.

2 Literature Survey

In 2018, Hussein et al. [12] proposed a system for detection of the lung cancer by computer-aided diagnosis (CAD) tools are often needed for fast and accurate

detection, characterization and risk assessment of different tumour from radiology images. And its accuracy obtained was 85%.

In 2019, Uc-ar et al. [13] recommended a new technique for lung cancer detection by using Laplacian and Gaussian filter along with the CNN architecture. They obtained accuracy by 72.97%.

In 2019 Preethi joon et al. projected respiratory cancer recognition by SVM classifier using fuzzy c & k-mean partition methodologies, and accuracy obtained is less than 90% [14].

Aggarwal et al. proposed a method for selection of the lung cancer by content-based medical retrieval, and they use different algorithms for detection. In this method, 86% precision is attained [15].

Vishal Patil et al. proposed a method for detection of lung cancer by using the pre-processing and segmentation. In this, they also did the feature extraction of the tumour and they got accuracy as 91% [16].

All of these methods achieve accuracy of less than or equal to 93%. As a result, we describe a novel method for detecting and classifying lung cancer disease that employs optimization to improve accuracy.

3   Methodology

See Fig. 1.

3.1   Pre-processing

Pre-processing is a technique involving the transformation of input data into understandable format. Pre-processing is aimed at improving the quality and optimizing the image quality and optimizing the image functionality further pre-processing. Normally, every image contains low-frequency noise, so the image has to undergo pre-processing to remove the noise [15]. As evident from Fig. 1, in pre-processing the stage involves is resize, and the collected input images are at different sizes. Therefore, we should establish the base size for all the input images. The base size of the image is 256×256. After resizing the image, we go for RGB to greyscale conversion for converting true colour image RGB to greyscale image.

Median Filter The median filter is used for enhancement in pre-processing. It's a filter that isn't linear. It dampens the sound of salt and pepper. The image's edges are preserved by the filter. In pictures, salt-and-pepper noise is an impulsive form of noise [15]. Errors in data transmission are the most common cause of noise. It appears as a smattering of white-and-black pixels that can substantially degrade the clarity of an image. It can be calculated by using mathematical expression.
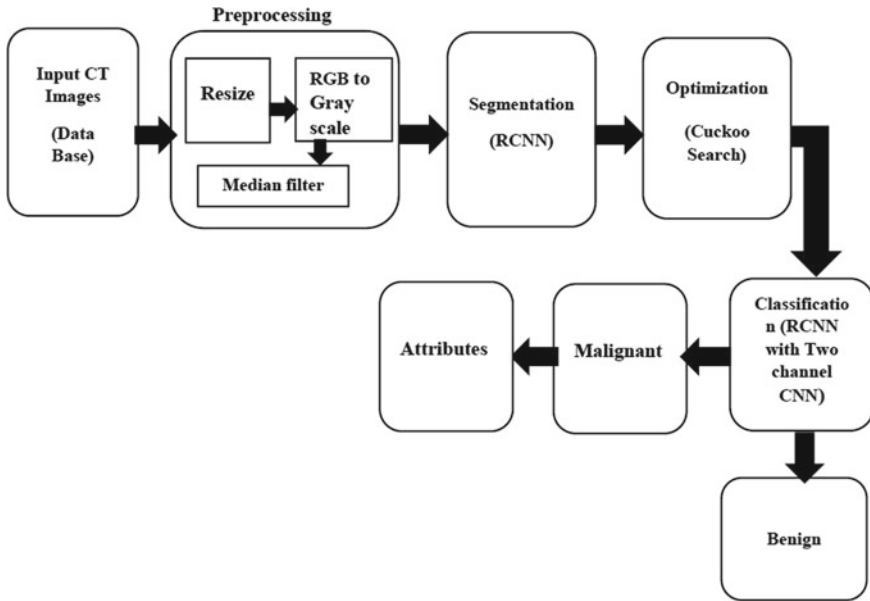
Fig. 1 Block diagram for proposed methodology using RCNN and two-channel CNN

$$M(g) = \sum_{k=1}^{n} |x_k - g|, \quad \text{where } g = \text{median}\{x_1, x_2, \ldots x_n\}$$

The reduction of noise in an image can be achieved by signal pre-processing. This filter replaces the through entry with the centre of the nearby entries to execute the image entry. The window, which slides in full signal or picture entry by entry, is known as the neighbours' pattern.

3.2 Segmentation with RCCN

Segmentation is a method of dividing an image into many parts or sections, usually based on the properties of the pixels in the image. It was strongly associated with visual objects [15]. The RCNN technique is used in this segmentation. R-original CNN's goal was to take an input image and generate a set of bounding boxes, each of which contained an item and the object's category (e.g. car or pedestrian). RCNN has recently been updated to include new computer vision tasks. RCNN starts by extracting regions of interest (ROI) from an input image, each ROI being a rectangle that could represent the image's boundaries [17]. It essentially comprises of a region proposal layer, a detection layer and CNN. The proposal layer divides the image into candidate regions, which are subsequently sent to CNN to be compared to

Fig. 2 RCNN
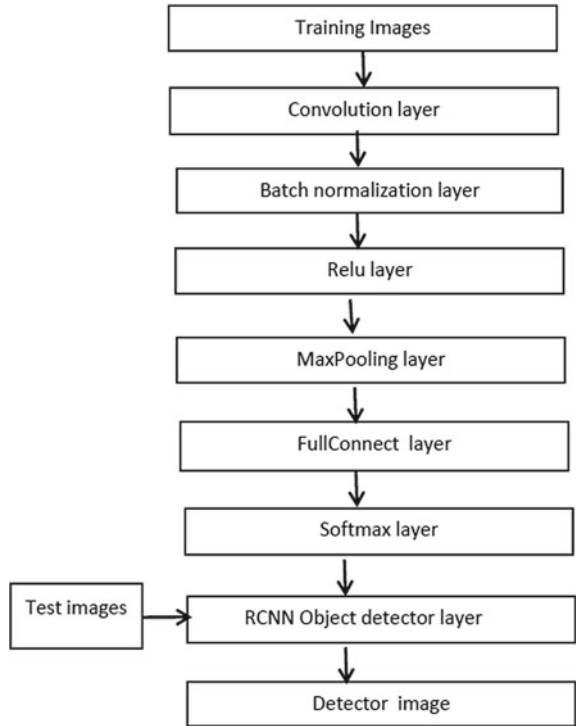
ground truth bounding boxes. The final convolution layer's feature vector is sent to the detection layer's fully convolutional layer. The final convolution layer's feature vector is fed into the fully connected layer, which is a classifier in this case. It determines whether or not the required objects are present. The RCCN is slow by nature, and it takes a long period to train it [13]. The processing steps in RCNN are shown in Fig. 2. As a result, we won't be able to retrieve the optimized values, so we'll use cuckoo search to optimize.

3.3 Optimization by Cuckoo Search

The practice of modifying a system to make some elements work more effectively or discovering alternative performance under given constraints as efficiently as possible by maximizing desired parameters and minimizing undesired parameters is referred to as optimization. Maximization is the process of attempting to achieve good results without spending a lot of money [14]. The cuckoo optimization algorithm (COA) is inspired by the lifestyles of birds in the cuckoo family [18]. The unusual lifestyle of these birds, as well as their characteristics in egg laying and breeding, inspired

the development of this new evolutionary optimization technique. Like other evolutionary techniques, the cuckoo optimization algorithm (COA) starts with a population. In diverse communities, there are two types of cuckoos: mature cuckoos and eggs. The cuckoo optimization algorithm is inspired by cuckoos' struggle to survive [19]. During the struggle for survival, some of the cuckoos and their eggs perished. Surviving cuckoo groups move to a more favourable environment and begin breeding and laying eggs. Cuckoos' struggles to survive should hopefully result in a situation where there is only one cuckoo society in which all cuckoos live [14]. Levy flight is a random walk in which the steps are determined by their lengths, which have a probability distribution, and the directions are random. Animals and insects both exhibit this random walk [20]. The following movement is determined by the current position [16]. The flowchart of cuckoo search optimization is shown in Fig. 3.

$$X_i(t + 1) = X_i t + \alpha \oplus \text{Levy}(\lambda) \qquad (1)$$

The step size is > 0 in this case. In the vast majority of circumstances, assume that is equal to one. The term "product" refers to entry-wise multiplication, often known as the "exclusive OR" operation. A random walk with random step size following a levy distribution is known as levy flight.

$$\text{Levy } u = t - \lambda 1 < \lambda \le 3 \qquad (2)$$

Algorithm for Cuckoo Search

- Objective function $q(x)$, $X = (X_1, X_2, X_3 \ldots X_f)$
- Provide an original population of n host nests;
- 222X $k(k = 1, \ldots, n)$
- While
- Get occasional cuckoo (say, k) Post energy level $q(X_k)$ and replace it's solution with levy flights;
- Access the energy level q_k
- Choose a nest at random among n (say, k):If $p_k < p_i$ then
- Substitute 1 by a new approach
- End if
- A fraction of q_a from the worst nests is abandoned and new ones are constructed;
- Keep the best solutions as nests;
- make a list of solutions/nests and assess the best current solutions;
- Pass the new best option H on to the next generation.
- End while.

Flowchart

See Fig. 3.
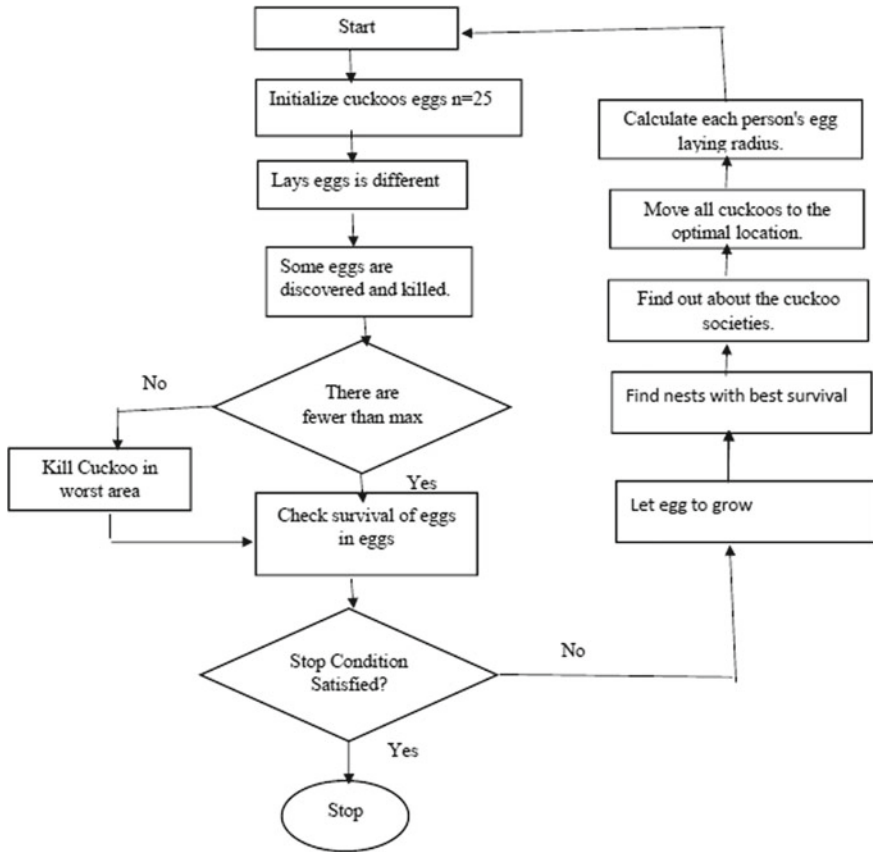
Fig. 3 Flowchart for cuckoo search optimization

3.4 Classification RCCN with Two-Channel CNN

In the classification procedure, the RCNN with two-channel CNN technique is used. We employ one channel CNN in the current technique. CNN kernels move in one direction in one channel and are employed in time series data. Because of the disadvantages of a single channel CNN, the suggested solution employs a two-channel CNN. Two-channel CNN is used in our proposed approach to categorize images from input photographs before sending them to an RCNN-based detection system, where the feature extraction layer of RCNN is substituted with two-channel CNN, which performs better. The two-channel CNN consists of two convolutional layers and a pooling layer, as shown in Fig. 4. The suggested method uses a two-channel CNN with RCNN to classify the images. The pictures are classed as benign or malignant [13].
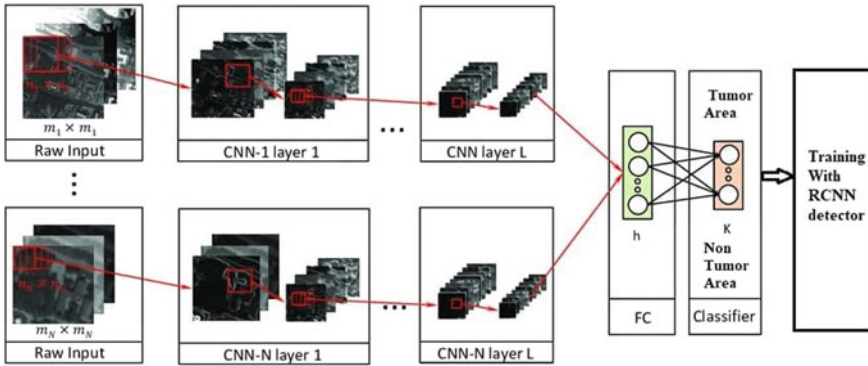
Fig. 4 Two-channel CNN with RCNN

Benign tumours are not cancer. There is a nuclear variation in size and form marginal. It is diploid and has low mitotic count, normal tissues. Representation of specialization differentiation is preserved and arranged.

A malignant neoplasm is a tumour made up of cancer cells. Nuclear variation in size and form ranges from minor to significant and is frequently varied. Ploidy ranges from low to high mitotic count, resulting in improper cellular division. In the classification procedure, the RCNN with two-channel CNN technique is used.

3.5 Attributes

Calculations of attributes area unit sometimes accustomed describe the metameric neoplasm image what quantity space is affected and a few of the parameters are: accuracy, sensitivity, specificity, peak signal-to-noise ratio (PSNR), entropy.

Accuracy Accuracy is defined as the percentage of correctly classified cases. It is calculated using the following mathematical formula:

$$\text{Accuracy} = (Tp + Tn)/(Tp + Fp + Fn + Tn)$$

Sensitivity Sensitivity measures the proportion of actual positive values that area unit properly known. It is calculated by exploitation mathematical expression:

$$\text{Sensitivity} = Tp/(TP + Fn)$$

Specificity Specificity is that the proportion of individuals while not the unwellness World Health Organization can have negative result. It is calculated by exploitation mathematical expression:

Fig. 5 Input image
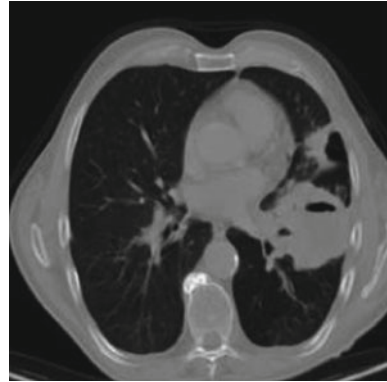
$$\text{Specificity} = Tn/(Tn + Fn)$$

Peak signal-to-noise magnitude relation (PSNR) The peak signal-to-noise magnitude relationship is a noise expression that may be determined using the following mathematical formula:

$$\text{Peak signal-to-noise ratio(PSNR)} = \text{ten} * \log 10(256^2/\text{mse})$$

4 Results and Analysis

CT images were acquired from public and private sources for this study. We employ RCCN segmentation for detection in this case and optimization by cuckoo search. In this, the input image, resized image and noisy image are in Figs. 5, 6 and 7, respectively. The median filter is shown in Fig. 8. Detection is performed through segmentation by RCNN and optimization through cuckoo search is shown in Fig. 9. Classification is performed by two-channel CNN with RCNN and classified as benign and malignant as shown in Fig. 10. Figures 11 and 12 depict the plot of images and the message displayed upon classification of the image, respectively. The statistical values obtained for different parameters have been shown in Table 1.

5 Conclusion

Lung cancer is one of the foremost dangerous diseases in the world. The detection of this disease at an early stage is kind of with the present methods of segmentation. CT images are pre-processed from the database, and it removes the noise through median

Fig. 6 Resized image
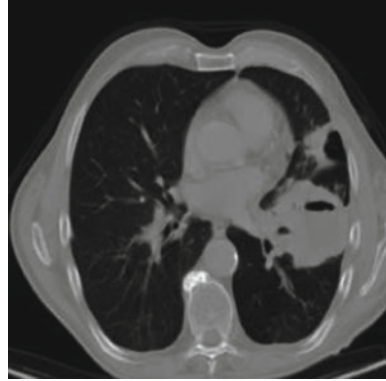
Fig. 7 Noisy image

Fig. 8 Filtered image
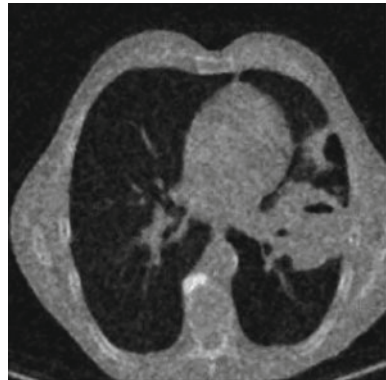
Fig. 9 Cuckoo search
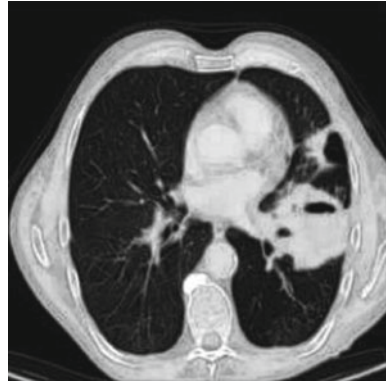
Fig. 10 Detected image
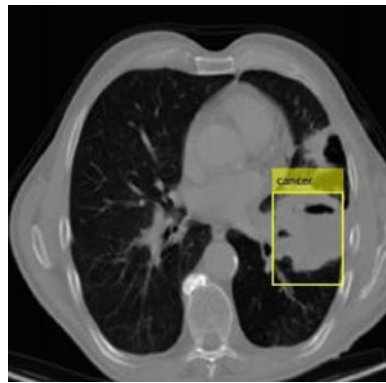
filter. Later the image is segmented by RCNN technique in conjunction with the improvement of cuckoo search, and additionally the feature extraction is performed and classified by using the RCCN with the two-channel CNN technique. The size and placement of illness are detected. Finally, the detected illness image classifies it's benign or malignant. Later parameters like accuracy, sensitivity, specificity, entropy, mean square error rate (MSE) are obtained. During this, we have a tendency to use the IoT for information transferring from MATLAB. The later parameters are transmitted through PhP server to mobile phones/PC.
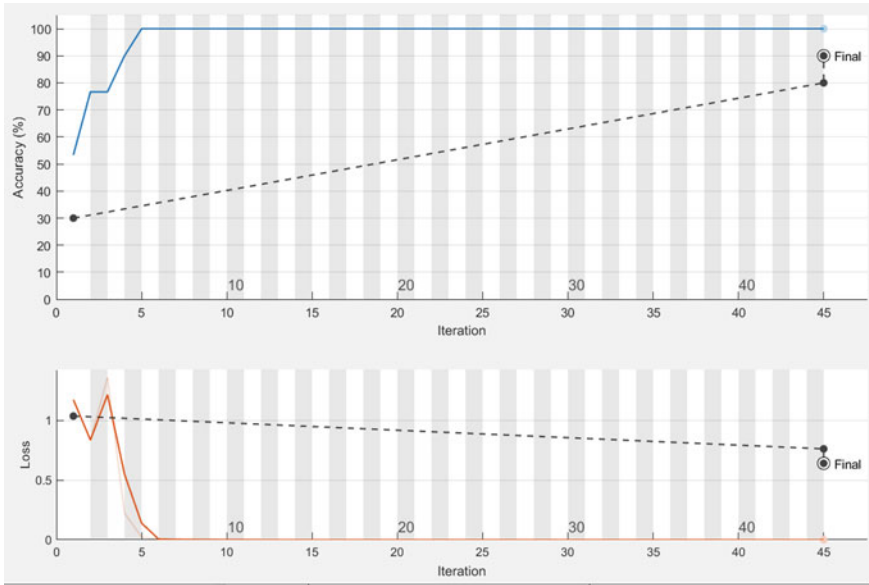
Fig. 11 Plot of images

Fig. 12 Classified image
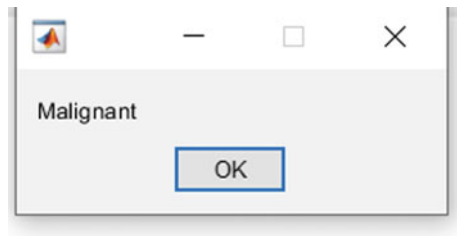
Table 1 Parameters values

Parameters	Statistical values
Accuracy	97.24
Precision	69.7948
Sensitivity	99.9946
Recall	99.9946
F measure	82.2089
Specificity	195.5012
AUC	99.9946

References

1. Perumal S, Velmurugan T (2018) Lung cancer detection and classification on CT scan images using enhanced artificial bee colony optimization. Int J Eng Technol 7(2.26):74–79
2. Bari M, Ahmed A, Sabir M, Naveed S (2019) Lung cancer detection using digital image processing techniques. Mehran Univ Res J Eng Technol 38(2):351–360. p-ISSN: 0254-7821, e-ISSN: 2413-7219. https://doi.org/10.22581/muet1982.1902.10
3. Rahane W, Magar Y, Dalvi H, Kalane A, Jondhale S (2018) Lung Cancer Detection Using Image Processing and Machine Learning healthcare. In: Proceeding of 2018 IEEE international conference on current trends toward converging technologies, Coimbatore, India
4. Zhanga S, Hanb F, Lianga Z, Tane J, Caoa W, Gaoa Y, Pomeroyc M, Ng K, Hou W (2019) An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets. Comput Med Imaging Graph 77
5. Liu H, Zhang S, Jiang X, Zhang T, Huang H, Ge F, Zhao L, Li X, Hu X, Han J et al (2019) The cerebral cortex is bisectionally segregated into two fundamentally different functional units of gyri and sulci. Cereb Cortex 29:4238–4252
6. Dolz J, Desrosiers C, Wang L, Yuan J, Shen D, Ayed IB (2020) Deep CNN ensembles and suggestive annotations for infant brain MRI segmentation. Comput Med Imaging Graph 79 [CrossRef]
7. Teramoto A, Tsukamoto T, Kiriyama Y, Fujita H (2017) Automated classification of lung cancer types from cytological images using deep convolutional neural networks. BioMed Res Int 2017
8. Trigueros DS, Meng L, Hartnett M (2018) Enhancing convolutional neural networks for face recognition with occlusion maps and batch triplet loss. Image Vis Comput 79:99–108
9. Giménez M, Palanca J, Botti V (2020) Semantic-based padding in convolutional neural networks for improving the performance in natural language processing. A case of study in sentiment analysis. Neurocomputing 378:315–323
10. Suresh S, Mohan S (2019) NROI based feature learning for automated tumor stage classification of pulmonary lung nodules using deep convolutional neural networks. J King Saud Univ Comput Inf Sci
11. Girshick R, Donahue J, Darrell T, Malik J Rich feature hierarchies for accurate object detection and semantic segmentation Tech report (v5). UC Berkeley
12. Hussein S, Kandel P, Bolan CW, Wallace MB (2019) Computer aided diagnosis (CAD) tools. IEEE, 2894349
13. Uc-ar et al (2019) Recommended a new technique for lung cancer detection by using Laplacian and Gaussian filter along with the CNN architecture
14. Uc-ar et al (2019) Recommended a detection model by Laplacian and Gaussian filter model with CNN architecture
15. Preethijoon et al (2019) Projected a respiratory cancer recognition strategy with the SVM classifier using fuzzy c & k-mean partition methodologies
16. Aggarwal P, Vig R, Sardana HK Semantic and content-based medical image retrieval for lung cancer diagnosis with the inclusion of expert knowledge and proven pathology. IEEE. https://doi.org/10.1109/ICIIP.2013.6707613
17. Kesav N, Jibukumar MG Efficient and low complex architecture for detection and classification of Brain Tumor using RCNN with Two CNN. https://doi.org/10.1016/j.jksuci.2021.05.008
18. Pentapalli1 VVG, Varma RK (2016) Cuckoo search optimization and its applications: a review. Int J Adv Res Comput Commun Eng 5(11). ISO 3297:2007 Certified
19. Venkatesh C, Sai Dhanusha K, Lakshmivara Prasad C, Areef S (2020) Deep learning based lung cancer detection in CT scans and secure data transmission. Int J Anal Exp Modal Anal XII(III)
20. Yang X-S, Deb S (2010) Engineering optimization by Cuckoo search. J Math Model Numer Optimisation 1(4)

Design of High Efficiency FIR Filters by Using Booth Multiplier and Data-Driven Clock Gating and Multibit Flip-Flops

P. Syamala Devi, D. Vishnu Priya, G. Shirisha, Venkata Tharun Reddy Gandham, and Siva Ram Mallela

Abstract In today's digital signal processing (DSP) applications, power optimization is one of the most significant design goals. The digital finite duration impulse response (FIR) filter is recognized as one of the most important components of DSP, and as a result, researchers have done numerous significant studies on the power refinement of the filters. FIR filters with conventional multipliers and adders are used through which the power consumption is increased. Data-driven clock gating (DDCG) is a technology that uses multibit flip-flops (MBFFs) to share a single clock enabling signal, optimize clock latency, manage clock skew, and improve routing source utilization. To achieve low-power consumption, these two low-power design strategies are integrated. The Xilinx ISE 14.7 software is used to accomplish low-power optimization.

Keywords FIR filter · DDCG · MBFF · Signal processing · Flip-flops

1 Introduction

1.1 Background

When it comes to building electrical gadgets like phones, power consumption is a major consideration. Static, dynamic, leakage, and short-circuit power are all sources of power in digital electronic circuits. Low static power is the primary advantage of CMOS VLSI circuits, yet dynamic power is the most powerful power source of all. The clock signal with the highest switching rate is the source of dynamic power usage. The finite impulse response (FIR) filter, on the other hand, is widely used as a major element for developing multiple digital signal processing (DSP) hardware

P. Syamala Devi (✉) · D. Vishnu Priya · G. Shirisha · V. T. R. Gandham · S. R. Mallela
Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: psd@aitsrajampet.ac.in

circuits due to its guaranteed linear phase and stability. These circuits are utilized in a variety of modern mobile computer and portable multimedia applications, including high-efficiency video coding (HEVC), channel equalization, voice processing, and software defined radio (SDR). This fact prompted designers to look for innovative ways to keep the FIR filter's power consumption low.

The FIR filters must be implemented in reconfigurable hardware in numerous applications, such as the SDR channelizer. By decreasing the filter coefficients without changing the order, the authors were able to reduce the FIR filter's power consumption. Alam and Gustafsson [1] used an approximation signal processing approach. Using add and shift operations, the filter's design is simplified in a variety of ways. Many strategies are utilized for low-power designs [2]. In [3], the technique used is called reduced dynamic signal representation. In [4], a reversible technique was used. In [5], a digital logic FIR filter is proposed. For FIR power minimization, a multibit flip-flop (MBFF) approach was presented in [6]. The data-driven clock gating (DDCG) technique was utilized to optimize power digital filters in [7]. In the last decade, several research using the clock gating technique for filter coefficients have been proposed. Data-driven clock gating has been implemented for digital filters.

1.2   Objectives

When clock pulses are not in use, the major goal of a Booth multiplier using clock gating technology is to reduce undesired switching activity. Clock gating is a commonly utilized technology that offers very effective solutions for lowering dynamic power dissipation. For processing, a clock gating module is employed. This necessitates a separate clock signal as well as an enable signal. Due to the use of separate clock and enable signals, each individual module consumes more power, increasing the area and power consumption. They are fed into an OR gate, which adds all the data and feeds it into a digital circuit. The circuit diagram, on the other hand, is smaller and employs fewer components, which reduces the amount of energy consumed by the components as well as the amount of space needed. Only a switching and one enable signal are required in the second circuit, which are both supplied into an AND gate.

The following are some examples of clock gating techniques used for dynamic power reduction:

- Clock gating based on gates
- Clock gating based on latches
- Clock gating based on flip-flops
- Clock gating based on synthesis
- Clock gating depending on data
- Clock gating that is automatically controlled
- Clock gating based on looking forward.

Data-driven clock gating is a flip-flop that can be made without the need for a latch, but it is used here to store the OR gate output for a complete cycle, reducing glitches. The data-driven flip-flop requires less power than the AGFF since both latches in the D FF are gated. The FF's clock signal can be halted in the following cycle by XORing an FF's output with the current data input, and that response will appear in the following cycle.

1.3 *Need and Importance*

Every circuit nowadays has to deal with the issue of facility use. To avoid this problem in the IC style for mobile devices and to reduce power dissipation in electronic circuits, numerous concepts were developed. The chip's power is depleted by an IC-style continuous system of temporal order parts such as flip-flops, latches, and clock networks. Three primary aspects play a critical part in IC style: area, power, and delay. In general, the higher-improvement strategy reduces facility consumption. Performance maximizing and power-minimization and style optimizations yielded facility and delay trade-offs. A number of researchers have altered clock gating approaches in a variety of ways. This study looks at misuse of power in register circuits, as well as clock gating approaches and multibit flip-flops. In general, the higher-improvement strategy reduces facility utilization. Performance maximization and power minimization style optimizations yielded facility and delay trade-offs.

2 **Existing Methods**

An Nth order FIR filter is used to execute N-point linear combining of the input pattern with input variables for an incoming data sample.

FPGA, on the other hand, comes at a cost in terms of performance, efficiency, and complexity as compared to ASICs [8]. The filter's generalized reconfigurable nature limits its performance improvement through algorithmic programme reformulation as provided in Fig. 1. In recent times, various designs have been proposed for this. A filter will be applied to the direct type (DF) or the backward direct type (BDT) within the direct type (DF) or the backward direct type (BDT) (TDF). The structure of the DF determines the TDF of the FIR filter. In terms of backward and direct types, Associate in Nursing FIR filter is equal. It is simple to show that the word length of every delay element is identical to the sign's word length in the direct form [9]. However, in the backward kind, every delay component has a lengthier word length than in the direct form; also, the lag parts want to postpone the merchandise or product addition. Although the backward structure reduces critical route latency, it uses a lot of hardware. The DF has one number + (M 1) adders within the basic path, whereas the TDF just has one multiplier + one adder. There has been a dramatic improvement in performance for big M [10, 11]. The TDF is preferable over the
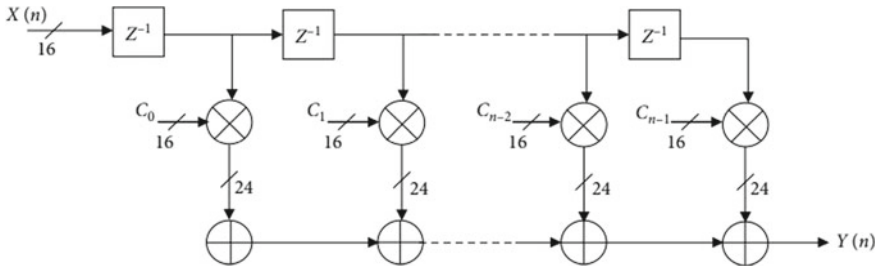
Fig. 1 FIR filter architecture

direct form in VLSI implementation because of its pipelined accumulation part. A TDF has two modules that create the filter output: multiple constant multiplication (MCM) and merchandise accumulation [12].

3 Proposed Methodology

The simplest way to make a digital filter is to combine the input signal with the digital filter's impulse response. As seen in Fig. 2, FIR filtering is characterized by a simple cross-linking process.

The input, filter output, and a parameter are represented as $x[n]$, $y[n]$, and $h[n]$, respectively. N [13] is used to represent the filter order. The multiplication is the most essential limitation in a FIR filter because it dictates the intended filter's performance. Modified Booth multipliers are commonly used for good performance. By processing three bits at a time, the Booth multiplier uses the Booth encoding algorithm to lessen the quantity of partial products during rewriting. This recoding approach is extensively used to generate intermediate results for big parallel multipliers that use the parallel encoding scheme.

Fig. 2 FIR filter
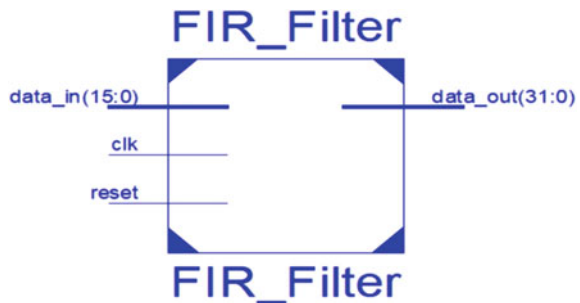
3.1 Using MATLAB

The input file is generated from MATLAB with four steps

1. Generate sine wave
2. Add noise
3. Convert real to integer
4. Convert integer to binary
5. This data is saved in signal.data file.

3.2 RTL Schematic Diagram

Many finite impulse response (FIR) filter designs have been created with a focus on small area, fast speed, and low energy consumption as shown in Fig. 3. It is clear that as the size of such FIR filters expands, so does the hardware cost of these filters. As a result of this discovery, a low-cost FIR filter with low-power and modest speed performance was developed. Simultaneous multiplier of distinct delay signals and their related filter coefficients is performed by the modified Booth multiplier module, followed by the accumulation of all results. An adder is also used for accumulation. For final addition, an area efficient adder, especially through the parallel prefix adder, is utilized to reduce the multiplier's area. Here, a modified Booth multiplier is built to improve the efficiency.
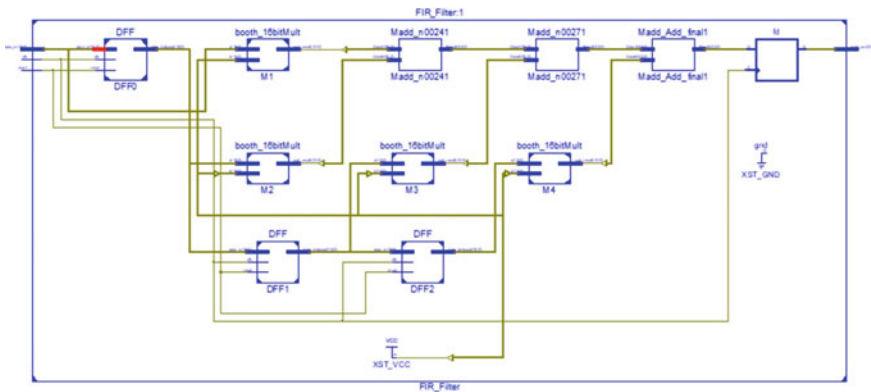
Fig. 3 FIR filter with booth multiplier
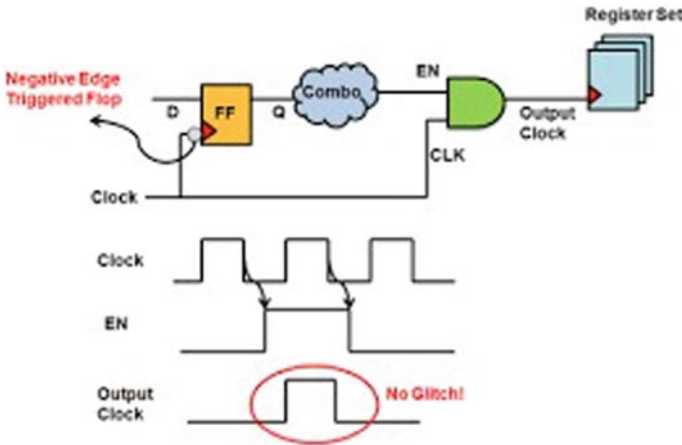
Fig. 4 Data-driven clock gating

3.3 Data-Driven Clock Gating

Clock gating is a technique for decreasing switched mode power consumption in clock signals as given by Fig. 4. When the current and next phases of the D flip-flop are examined, it is discovered that the D flip-flop produces the same result as the output when two consecutive inputs are identical. The latch uses clock power even though the sources do not vary by one clock to the next.

Driven by data clock gating circuitry is a popular methodology for decreasing dynamic power waste which is synchronous circuits by eliminating the clock signal when the circuit does not require it. The term "data driven" refers to the fact that data is provided as input to the clock. We can use the AND gate to access this clock.

3.4 Implementation Tool

The tool that we are most likely to utilize throughout this project is the Xilinx tool, which is generally used for circuit synthesis and design, whilst ISIM, or the ModelSim logic machine, is used for system level testing. As is customary in the business electronic design automation industry, Xilinx ISE is tightly connected to the design of Xilinx's own chip and cannot be used with FPGA products from other vendors.

System-level testing can be done with ISIM or the ModelSim logic machine, and such tests should be written in alpha-lipoprotein languages. Simulated signalling waveforms or monitors that monitor and verify the outputs of the device under test are examples of test bench programmes.

ModelSim or ISIM can also be used to do the following types of simulations:
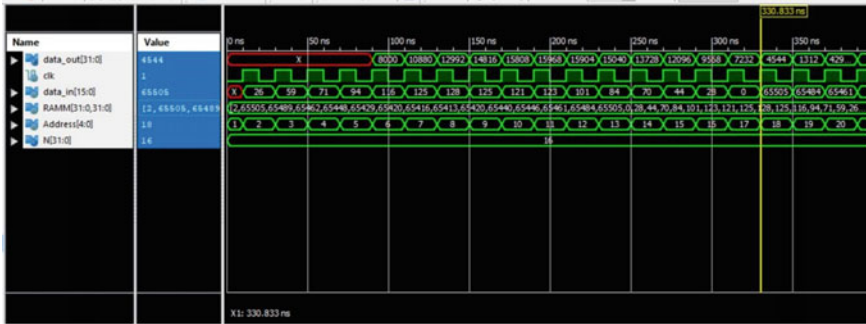
Fig. 5 Timing analysis diagram

- Logical testing to ensure that the module generates the anticipated outcomes.
- Behavioural verification is used to check for logical and temporal arrangement issues.
- Post-place and route simulation to ensure proper behaviour when placing the module at intervals in the FPGA's reconfigurable logic.

4 Simulation Results

The simulation result for proposed FIR filter as shown in Fig. 5 is represented by means of a timing diagram. The inputs along with the clock input are available at the right side of the figure, whereas timing diagram indicates the delay in seconds.

5 Conclusion

We suggested the FIR filter with Booth multiplier in this work, which results in reduced power consumption. The suggested Booth multiplier works by decreasing partial products, resulting in excellent precision. In this manner, the computationally costly component of the multiplication was removed, resulting in increased speed and energy consumption at the cost of the tiny inaccuracy. This approach could be used to solve both signed and unsigned multiplications. The recommended Booth multiplier has a route delay of 1.933 ns and can save up to 9.6% of power.

References

1. Alam SA, Gustafsson O (2014) Finite word length design. In the logarithmic number system, linear-phase FIR filters. VLSI Des 2014:14. Article ID 217495
2. Proakis J, Manolakis D (2006) Digital signal processing: principle, algorithms and applications, 4th edn. Prentice-Hall, Upper Saddle River, NJ, USA
3. Chen K-H, Chiueh T-D (2006) A low-power digit-based reconfigurable FIR filter. IEEE Trans Circuits Syst II Express Briefs 53(8):617–621
4. Hwang S, Han G, Kang S, Kim J (2004) New distributed arithmetic algorithm for low-power FIR filter implementation. IEEE Signal Process Lett 11(5):463–466
5. Carloni LP, De Bernardinis F, Pinello C, Sangiovanni-Vincentelli AL, Sgroi M (2005) Platform-based design for embedded systems. In: Zurawski R (ed) The embedded systems handbook. CRC Press, Boca Raton, FL, USA
6. Wimer S, Koren I (2014) Design flow for flip-flop grouping in data-driven clock gating. IEEE Trans VLSI Syst 771–779
7. Mamatha B, Ramachandram VVSVS (2012) Design and implementation of 120 order filter based on FPGA. Int J Eng Sci Technol 90–91
8. Mahesh R, Vinod AP (2010) New reconfigurable architectures for implementing FIR filters with low complexity. IEEE Trans Comput Aided Des Integr Circuits Syst 29(2):275–288
9. Nagaraju CH, Durga TK (2015) Implementation of carry select adder with area-delay-power and efficiency. Int J Sci Eng Technol Res 4(56):11916–11920
10. Johansson K, Gustafsson O, Wanhammar L (2007) Bit-level optimization of shift-and-add based FIR filters. In: Proceedimgs of the 2007 14th IEEE international conference on electronics, circuits and systems, vol 3, pp 713–716, Marrakech, Morocco
11. Nagaraju CH, Sharma AK, Subramanyam MV (2018) Reduction of PAPR in MIMO-OFDM using adaptive SLM and PTS Technique. Helix Vol 8(1):3016–3022
12. Gamatié A, Beux S, Piel E et al (2011) A model-driven design framework for massively parallel embedded systems. ACM Trans Embed Comput Syst 10(4)
13. Seok-Jae L, Choi J-W, Kim SW, Park J (2011) A reconfigurable FIR filter architecture to trade off filter performance for dynamic power consumption. IEEE Trans VLSI Syst 19(12):2221–2228

Detecting the Clouds and Determining the Weather Condition and Coverage Area of Cloud Simultaneously Using CNN

M. Venkata Dasu, U. Palakonda Rayudu, T. Muni Bhargav, P. Pavani, M. Indivar, and N. Sai Madhumitha

Abstract We need to consider certain elements of the climate system to forecast the climate; one of it is the role of clouds in evaluating the climate's sensitivity to change. Here, we will determine the area covered by cloud and the weather condition at specific time. Before performing this, we will detect the clouded part from satellite image using pre-trained U-Net Layers. Later, cloud coverage area and weather will be performed using CNN techniques. Experiments have shown that our suggested framework is capable of detecting and showing the cloud coverage region while also displaying the current weather conditions.

Keywords Atmosphere · Weather · Cloud identification · CNN

1 Introduction

The layer of atmosphere in which all clouds develop is about the same thickness as the leather cover of a softball when compared to the size of the earth. Water, the most ubiquitous and remarkable of all things, exists in a plethora of ever-changing forms within this delicate layer [1–3]. Water, unlike most other substances on the planet, can exist in all three phases—gaseous, liquid, and solid—within the small range of atmospheric conditions found on the planet. Clouds can emerge and disappear at any time, and "precipitate," pelting us with rain and snow as a result of this "versatility." While these occurrences may seem mundane to us, they are nothing short of extraordinary from a cosmic perspective [4–6]. And there's still a lot of unknowns. For example, scientists have no idea how ice crystals grow in clouds, and many clouds have significantly more than they expected. The sky itself may be a thrill, offering something fresh every day if we just look.

M. Venkata Dasu (✉) · U. Palakonda Rayudu · T. Muni Bhargav · P. Pavani · M. Indivar · N. Sai Madhumitha
Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: dassmarri@gmail.com

1.1   Objectives

The main goal of this project is to determine the area covered by cloud and the weather condition at specific time. Before performing this, we will detect the clouded part from satellite image using pre-trained U-Net Layers. Later cloud coverage area and weather will be performed using CNN techniques. Experiments have shown that our suggested framework is capable of detecting and showing the cloud coverage region while also displaying the current weather conditions.

1.2   Need and Importance

Cloud identification in satellite photography has a variety of uses in weather and climate research. Satellite images are important for monitoring changes in land covers such as forests, towns, agriculture, and coastal areas; however, clouds provide a challenge in most satellite imaging-based applications [7]. For satellite photos, precise detection is critical. Cloud pixels have a high rate of misclassification due to a lack of contrast between cloud edges and the land or sea background is a fundamental difficulty for cloud identification techniques [8]. Based on the percentage of the sky was obscured by opaque (not transparent) clouds, and the sky was overcast describes the predominant/average sky state [9]. If the sky condition is inferred from the precipitation forecast and there is a high probability of precipitation (60% or greater) for the majority of the forecast period, the sky condition may be omitted [10].

2   Literature Review

P. Dai, H. Zhang, L. Zhang, and H. Shen: In this research, a novel deep learning-based spatiotemporal fusion model is suggested for handling the large spatial resolution gap and nonlinear mapping between the high-spatial resolution (HSR) picture and the matching high-temporal resolution (HTR) image at the same imaging time. A two-layer fusion approach is used due to the large spatial resolution gap. The CNN model is used in each layer to take advantage of the nonlinear mapping between the HSR and HTR pictures to reconstruct the high-spatial and high-temporal (HSHT) resolution images. Landsat data is utilized to represent high geographic resolution photos in the experiment, whereas MODIS data is used to represent low spatial resolution images [11].

Drawbacks

- Calculating histogram equalization does not require a lot of computing power. This approach has the drawback of being indiscriminate. It may improve background noise contrast while diminishing useable pixel/signal.

- The SVM approach fails to work with large data sets. SVM does not work well when there is greater noise in the data set. If the number of characteristics for each data point is more than the number of training data samples, the SVM will perform poorly.

3 Methodology

3.1 Existing Method

For detection of clouds, some existing methods used histogram equalization techniques. They got the histogram of image by the command "imhist". And, they improved the intensity value over the full image by "histeq" command. Along with these steps, some preprocessing steps are performed such as image resizing, image rotating, and image addition. For determining/classifying the weather condition, methods like k-nearest neighbor and support vector machine (SVM) are used [12].

Histogram equalization is a computer-aided image processing technique for enhancing image contrast. It does so by effectively spreading out the most common intensity values, i.e., expanding out the image's intensity range. When the useful data is represented by close contrast values, this method frequently boosts the global contrast of images. This enables locations with low local contrast to obtain a boost in contrast. The number of pixels in each sort of color component is represented by a color histogram of a picture. Histogram equalization can't be applied independently to the image's Red, Green, and Blue components because it changes the color balance dramatically. However, if the image is first transformed to another color space, such as HSL/HSV, the technique can be applied to the luminance or value channel without changing the image's hue or saturation.

3.2 Proposed Method

The main theme of our project is to determine the area covered by cloud and the weather condition at specific time as shown in Fig. 1. Before performing this, we will detect the clouded part from satellite image using pre-trained U-Net Layers. Later cloud coverage area and weather will be performed using CNN techniques.

U-Net

A semantic segmentation system is the U-Net architecture as shown in Fig. 2. A contract path and an expansion path are both available. The contracting pathway is built using a standard convolutional network design. The name U-Net comes from the architecture, which resembles the letter U when visualized, as shown in the diagram below.
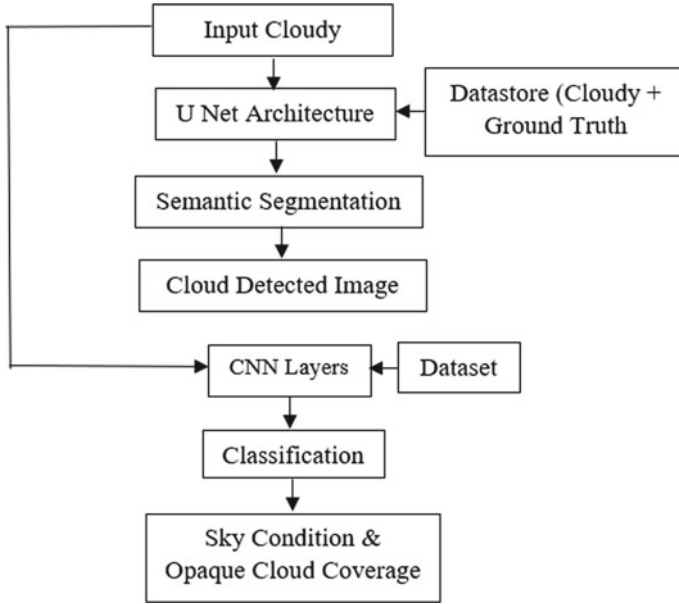
Fig. 1 Proposed method block diagram
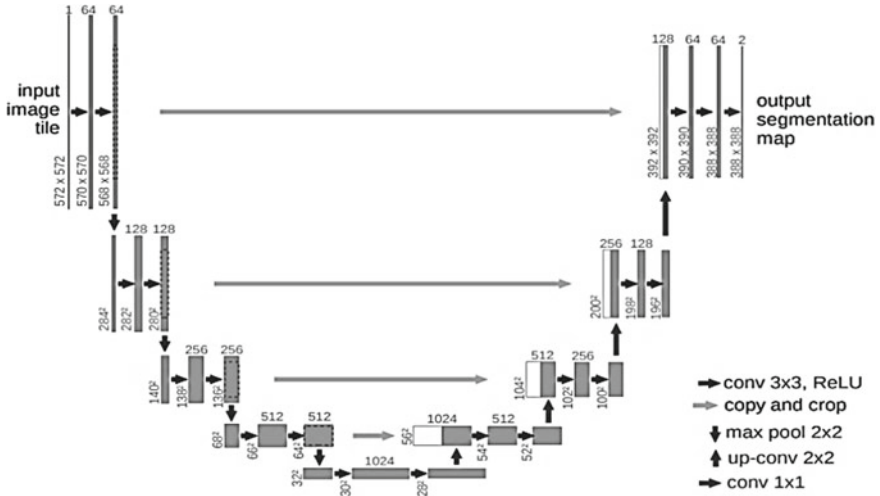
Fig. 2 Visualization of U-Net architecture

Semantic segmentation refers to the process of assigning a label to any pixel in a picture. This is in sharp contrast to categorization, which assigns a single label to the entire image. In semantic segmentation, many items of the same class are seen as a single topic.

CNN

Artificial neural networks perform exceptionally well in machine learning. Artificial neural networks are used for a wide range of classification tasks, including image, audio, and word classification. For example, we employ recurrent neural networks, more precisely an LSTM, to forecast the succession of words, and convolution neural networks for image categorization. In this part, we'll put together essential CNN building blocks. A convolutional neural network can have one or more convolutional layers. The number of convolutional layers utilized is determined on the volume and complexity of the data.

Let's review some neural network fundamentals before getting into the convolution neural network. There are three types of neurons in a standard neural network layers.

Image Input Layer: Use image input layer to create an image input layer. An image input layer sends images to a network and normalizes the data. Use the input Size option to specify the image size. An image's size is determined by the image's height, width, and number of color channels. The number of channels in a grayscale image is one, while it is three in a color image.

Convolutional Layer

Slider convolutional filters are applied to the input by a 2D convolutional layer. Convolution2dLayer is used to create a 2D convolutional layer. The convolutional layer is made up of several components. Filters and Stride are convolutional layers made up of neurons that link to subregions of the input pictures or the previous layer's outputs. While scanning a picture, the layer learns the features localized by these regions. The filter Size input argument can be used to specify the size of these sections when using the convolution2dLayer function to create a layer.

Software Requirements

For the simulation of the proposed method, MATLAB R2020a software is utilized. A special monitor, designated the MATLAB workspace, emerges whenever you start MATLAB.

4 Results

The input original image as shown in Fig. 3 is considered for simulation using the proposed model. The resulted image as shown in Fig. 4 indicates the segmentation of different regions, whereas Fig. 5 depicts the thresholded output which indicates

the detected cloud. Weather condition and its corresponding cloud coverage are displayed as message box at the end of simulation as shown in Fig. 6.

Fig. 3 Input image

Cloudy Image

Fig. 4 Segmented different regions

Segmented Different Regions
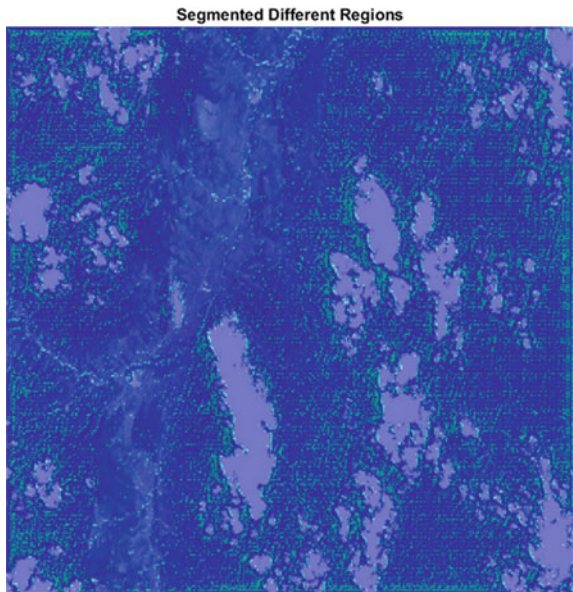
Fig. 5 Detected cloud image
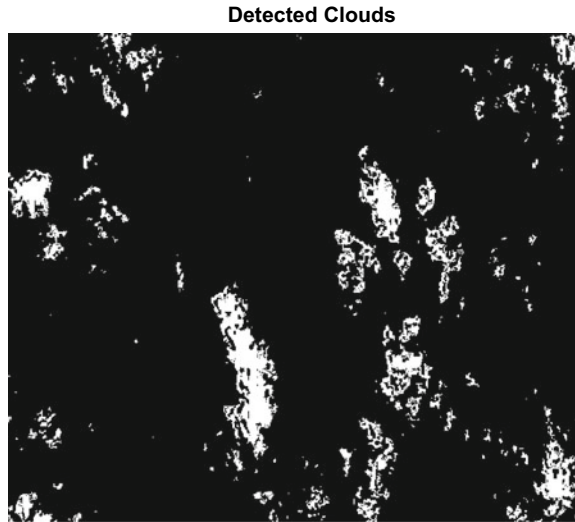
Detected Clouds

Fig. 6 Weather condition and cloud coverage

Sky Condition: Sunny / Clear
Opaque Cloud Coverage: 10%-30%

OK
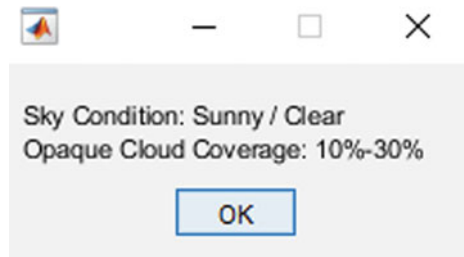
5 Conclusion

Here, we determined the area covered by cloud and the weather condition at specific time. Before performing this, we have detected the clouded part from satellite image using pre-trained U-Net Layers. Later, cloud coverage area and weather have performed using CNN techniques. Experiments have shown that our suggested framework can detect and display the coverage area of clouds while also displaying the weather state and gives better results when compared with existing works.

References

1. Nagaraju CH, Sharma AK, Subramanyam MV (2018) Reduction of PAPR in MIMO-OFDM using adaptive SLM and PTS technique. Int J Pure Appl Math 118(17):355–373
2. Tian B, Shaikh MA, Azimi-Sadjadi MR (1999) A study of cloud classification with neural networks using spectral and textural features. IEEE Trans Neural Netw

3. Yuan K, Meng G, Cheng D, Bai J, Xiang S, Pan C (2017) Efficient cloud detection in remote sensing images using edge-aware segmentation network and easy-to-hard training strategy. In: Proceedings of the IEEE international conference on image process (ICIP), Sept 2017, pp 61–65

4. Dai P, Zhang H, Zhang L, Shen H (2018) A remote sensing spatiotemporal fusion model of landsat and modis data via deep learning. In: Proceeding of the IEEE international geoscience and remote sensing symposium, July 2018, pp 7030–7033

5. Ji S, Wei S, Lu M (2019) Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. IEEE Trans Geosci Remote Sens 57(1):574–586

6. Braaten JD, Cohen WB, Yang Z (2015) Automated cloud and cloud shadow identification in landsat MSS imagery for temperate ecosystems. Remote Sens Environ 169:128–138

7. Fisher A (2014) Cloud and cloud-shadow detection in SPOT5 HRG imagery with automated morphological feature extraction. Remote Sens 6(1):776–800

8. Ackerman SA, Holz RE, Frey R, Eloranta EW, Maddux BC, McGill M (2008) Cloud detection with MODIS. Part II: validation. J Atmos Ocean Technol 25(7):1073–1086

9. Mateo-Garcia G, Gomez-Chova L, Camps-Valls G (2017) Convolutional neural networks for multispectral image cloud masking. In: Proceeding of the IEEE international geoscience and remote sensing symposium (IGARSS), July 2017, pp 2255–2258

10. Li Z, Shen H, Li H, Xia G, Gamba P, Zhang L (2017) Multifeatured combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery. Remote Sens Environ 191:342–358

11. Zhu Z, Woodcock CE (2012) Object-based cloud and cloud shadow detection in landsat imagery. Remote Sens Environ 118:83–94

12. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. SpringerPlus 5(507). https://doi.org/10.1186/s40064-016-2152-2

Traffic Prioritization in Wired and Wireless Networks

Md. Gulzar, Mohammed Azhar, S. Kiran Kumar, and B. Mamatha

Abstract Traffic prioritization is an important task in networks. Because not all communications are the same, some communications may neglect the delays but some communications are delay critical. If congestion occurs at the intermediate devices the packets will be dropped or communication will be delayed. Delay can be tolerated for some applications but not for all. If we prioritize the traffic such that delay critical communications get high priority and the communications which are not delay critical get least priority, then even when congestion occurs the least priority packets will be dropped or delayed and bandwidth will be given to higher priority traffic. There are different methods provided for prioritizing the traffic. In this paper, we will discuss the available techniques of traffic prioritizing in wired and wireless networks.

Keyword DiffServ · Priority · QoF · DCF · PCF · Hybrid coordinator

1 Introduction

In the present era we can see two different networks named as wired networks and wireless networks. In wired networks we will have cables connected on outgoing ports on which data will flow to the destination. Since the destination may not be in same network the data packets may flow through different intermediate devices like switches, routers and firewalls to reach the destination. These intermediate devices will have many incoming ports and outgoing ports. And each port will have queues to hold the data packets before they are processed at the intermediate devices if the network is congested, then the queues at incoming ports will be overflowed and packets needs to be dropped. This causes the sending station to retransmit the packets which are dropped which will cause the congestion situation to get worst. If we prioritize the packets based on type of application the least priority packets will be dropped in case of congestion. With this efficiency of network increases. In this

Md. Gulzar (✉) · M. Azhar · S. Kiran Kumar · B. Mamatha
CMR Engineering College, Medchal, Hyderabad, India
e-mail: md.gulzar@cmrec.ac.in

335

paper we will brief the available methods to prioritize the traffic in both wired and wireless networks.

2 Traffic Prioritization in Wired Networks

Traffic prioritization is a concept of Quality of Service (QoS). QoS enables network administrators to provide minimum bandwidth for less time critical applications and maximum bandwidth for real-time traffic like voice, video where delay is not tolerated. QoS parameter will be configured on switches and routers. Generally the traffic generated by stations is characterized and marked in to various classes. And the intermediate devices like routers and switches will have queues with Class of Service (CoS) category. When packet arrives at the device it will be moved to appropriate queue based on its Class of Service [1].

Network traffic can be classified into various classes which includes

a. Explicit 802.1p
b. DSCP marking
c. VLAN/switch port-based grouping.

2.1 Explicit 802.1p

It is also called as IEEE P802.1p which is a subset of IEEE 802.1D, which is later included into IEEE 802.1Q in 2014. It provides a method for implementing QoS at media access control (MAC) level. In 802.1p QoS is implemented as Class of Service (CoS), which includes 3 bits known as Priority Code Point (PCP). It defines 8 priority values from 0 to 8 [1].

Eight different Classes of Services are mentioned in PCP field of IEEE 802.1Q header which is added to the frame at data link layer. The IEEE has recommended following values as shown in Table 1 to each priority level but these are only the guidelines by IEEE, and the network administrators can define their own levels for implementations.

2.2 DSCP Marking

Differentiated Services (DiffServ) is a network layer (layer 3) QoS mechanism. In IPv4 header it was Type of Service (TOS). In IPv6 it was changed to Differentiated Services field (DS field). It uses a 6-bit Differentiated Services Code Point (DSCP) inside DS field of 8 bits. It supports 64 classes of network traffic. It classifies and marks packets to a specific class. DiffServ uses DiffServ-aware routers. DiffServ places each data packet into one of the traffic classes. Each traffic class will have

Table 1 Priority values used in explicit 802.1p

PCP value	Priority	Acronym	Traffic types
1	0 (lowest)	BK	Background
0	1 (default)	BE	Best effort
2	2	EE	Excellent effort
3	3	CA	Critical applications
4	4	VI	Video, < 100 ms latency and jitter
5	5	VO	Voice, < 10 ms latency and jitter
6	6	IC	Internet control
7	7 (highest)	NC	Network control

priority value. Packets are handled at routers based on their priority values. The edge routers at edge of the network will do classification and policing [3]. A traffic classifier inspects many parameters of incoming packets such as source address, destination address or traffic type and assign to specific traffic class.

The edge routers will commonly have following behaviors:

- **Default Forwarding (DF) PHB** which is used for best-effort traffic.
- **Expedited Forwarding (EF) PHB** which is used for low-loss, low-latency traffic.
- **Assured Forwarding (AF) PHB** which gives assurance of delivery under certain conditions.
- **Class selector PHBs**: these are backward compatible with precedence field of IP.
- The various priority values and classes suggested are as follows in Table 2.

These are only the suggested guidelines by a specific vendor of DSCP. The administrator can change the values according to implementation.

Table 2 Priority values in DSCP marking

S. No.	Traffic type	Priority
1	IP routing	48
2	Video conferencing	34
3	Streaming video	32
4	Mission critical	26
5	Call signaling	24
6	Transactional data	18
7	Network management	16
8	Bulk data	10
9	Scavenger	8
10	Best effort	0

2.3   VLAN/Switch Port-Based Grouping

In Ethernet network traffic management is done by using per-port flow control where we will have two directions of data flow. One is upstream; other one is downstream. The upstream is from source to destination, and downstream is in reverse direction. When congestion occurs in network an overwhelmed switch notifies about the congestion to ports of downstream switch to stop the transmission for some period of time. The disadvantage of this method is that the downstream switch stops all traffic out of corresponding outgoing port even when there is no congestion from other sources.

This concept is shown in below figure. The source of congestion is first downstream switch A, which causes congestion at core switch through aggregation switch. Core switch stops receiving frames from aggregation switch even when there is no congestion from second downstream switch B.

Another solution to above problem is to assign priority to each downstream switch. This is called as priority-based flow control mechanism which was proposed in IEEE 802.3 standard. When any downstream switch causes congestion on core switch it will identify the priority of that downstream switch and makes it to stop the transmission for some amount of time, while accepting traffic from other downstream switches which do not cause congestion [2].

As in Fig. 1, if priority of downstream switch A which was causing congestion is P(10), and other downstream switch B, whose priority is P(12). The core switch identifies the downstream switch A and informs it to stop the transmission while receiving traffic from downstream switch B.

3   Traffic Prioritization in Wireless Networks

Wireless Multimedia Extensions (WME) known as Wi-Fi Multimedia is a part of IEEE 802.11e MAC Enhancements for Quality of Service (QoS). This includes 802.11 extensions for improving application performance of WLANs that carry multimedia data. Then 802.11e contains two methods for prioritizing network traffic [1].

They are:

a. Enhanced distribution channel access (EDCA).
b. Hybrid coordination function (HCF).

3.1   Enhanced Distribution Channel Access (EDCA)

Wireless networks consist of two types of channel access, i.e., contention-based channel access and controlled channel access. Contention-based channel access is
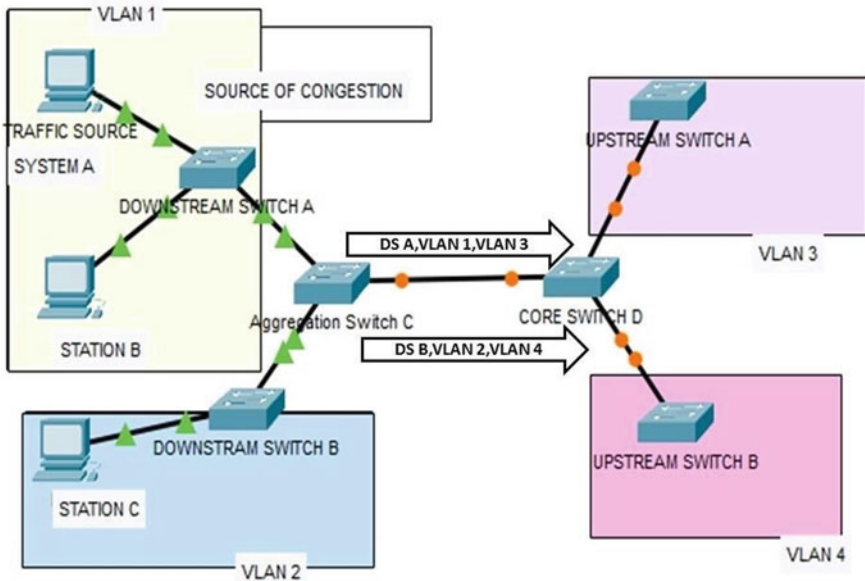
Fig. 1 Example for switch port-based grouping

also known as enhanced distribution channel access. The stations will access the channel in a specified limit known as Transmission Opportunity (TXOP). TXOP is defined by starting time and duration. The traffic is prioritized in EDCA by using access categories (ACs) and back off entities. ACs are prioritized according to AC-specific contention parameters. Depending upon traffic in selected access category contention window (CW) values including CWmin and CWmax are taken. And CWmax value is used for categories with higher traffic. EDCA has four access categories AC_VO (voice), AC_VI (video), AC_BE (best effort), and AC_BK (background) [1]. The contention window values are calculated as follows in Table 3.

The priority levels for each category and type of traffic are shown in Table 4.

Table 3 Formulas for values of contention window

AC	CWmin	CWmax
Background (AC_BK)	aCWmin	aCWmax
Best effort	aCWmin	aCWmax
Video (AC_VI)	(aCWmin + 1)/2 − 1	aCWmin
Voice	(aCWmin + 1)/4 − 1	(aCWmin + 1)/2 − 1

aCWmin = 15 and aCWmax = 1023 are typically used.

Table 4 Priority values in EDCA

Priority	Priority Code Point (PCP)	Access category (AC)	Traffic type
Lowest	1	AC_BK	Background
	2	AC_BK	Background
	0	AC_BE	Best effort
	3	AC_BE	Best effort
	4	AC_VI	Video
	5	AC_VI	Video
	6	AC_VO	Voice
Highest	7	AC_VO	Voice

3.2 HCF-Controlled Channel Access

Distributed coordination function (DCF)

Distributed coordination function (DCF) is a medium access technique which uses carrier sense multiple access with collision avoidance (CSMA/CA) which works on the principle of "Listen before Talk" with additional binary exponential back off.

Point Coordination Function (PCF)

IEEE 802.11 defines an access method known as point coordination function (PCF) which uses access point as network coordinator which manages channel access.

Hybrid coordination Function (HCF)

This is the combination of DCF and PCF. The controlled channel access of HCF is known as HCF-controlled channel access (HCCA). This method is same like EDCA allowing highest priority medium access to hybrid coordinator (HC) which is an access point during Contention Free Period (CFP) and Contention Period (CP) [2].

During CP this method works same like EDCA. HC works as central controller and polls the stations to send data frames. Each station sends its traffic needs periodically to HC. Based on this information, HCCA grants TXOP to the requesting station. The time during HC controls the access is known as controlled contention [6].

4   Reservation Protocols

In these protocols stations willing to transmit data will do broadcast to themselves before actual transmission. These protocols work in the medium access control (MAC) layer and transport layer. These protocols use a contention period prior to transmission. In the contention period, each station broadcasts its desire for transmission. After broadcasting each station gets the desired network resources. All

possibilities of collisions are eliminated because each station has complete knowledge about other stations regarding whether they want to transmit or not before actual transmission.

4.1 Bit-Map Protocol

In bit-map protocol we have contention period, which is used to decide the order of reservations. The contention period has n time slots, where n is the number of stations. Each station broadcasts its wish to transmit by placing 1 bit in its contention time slot. For example if there are 8 stations and Station 2 and Station 4 are willing to transmit, they first transmit their willingness by keeping a bit 1 in their contention time slots as shown in Fig. 2 [4].

As the final contention frame shows every station knows that Stations 2 and 4 are willing to transfer. And transmission happens according to the order of time slot in
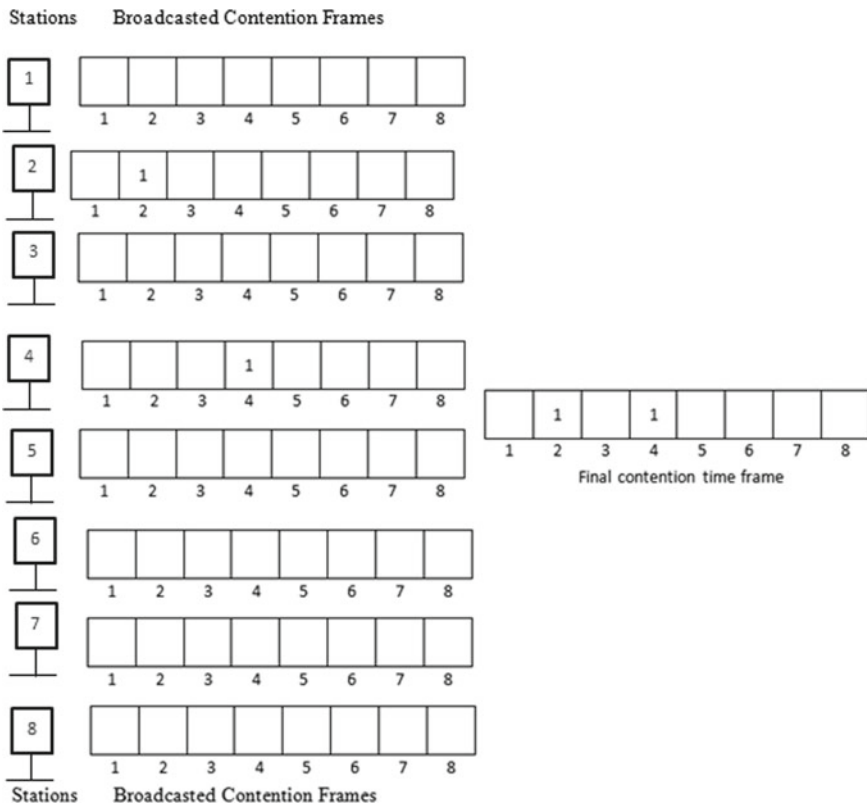
Fig. 2 Bit-map protocol

contention frame. That is first Station 2 transmits then Station 4. After contention frame time passes and data has been transmitted by station new contention time frame will be generated.

4.2   Resource Reservation Protocol (RSVP)

This protocol is used only for reservation on channels but not for actual data transmission. The resources which need to be reserved to avoid congestion or collision are bandwidth, buffer space and CPU cycles. If these are reserved on intermediate devices like routers, they can decide to accept or reject upcoming flow requests based on availability of resources. Every sender needs to specify its request for resources to routers and intermediate devices. These are called as flow specifications. RSVP is used to make the reservations of other protocols that are used to send the data. RSVP allows multiple senders to send data to multiple receivers, i.e., multicasting. In RSVP receivers will make the reservations not the senders. RSVP uses two types of messages: they are Path Messages and RESV Messages [5].

Path Messages. As we know in RSVP receiver will make reservation but it will not know the path in which sender will send packets to it. To solve this problem RSVP uses Path Messages. Path Messages travel from sender to receiver via multicast path. When Path diverges a new Path Message will be created as shown in Fig. 3.

RESV Message. After receiving a Path Message the receiver will send RESV Message toward the senders and makes reservations at the routers which support RSVP. If router does not support RSVP it provides best-effort service (Fig. 4).

Reservations are not provided for each flow, and the reservations are merged. That is if a flow needs 2 Mbps of bandwidth and another flow needs 1 Mbps of bandwidth.
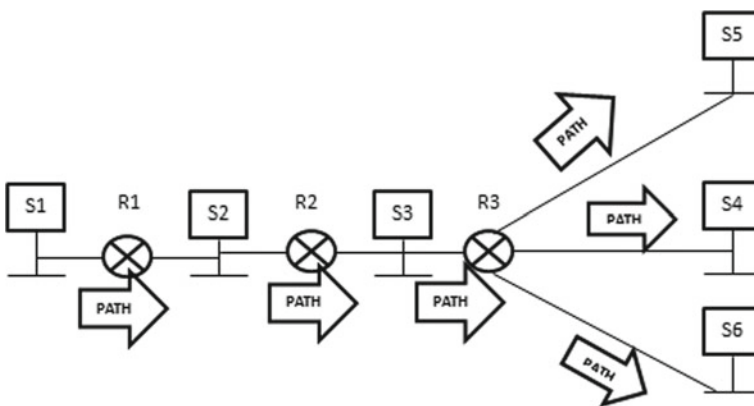
Fig. 3 RSVP path messages
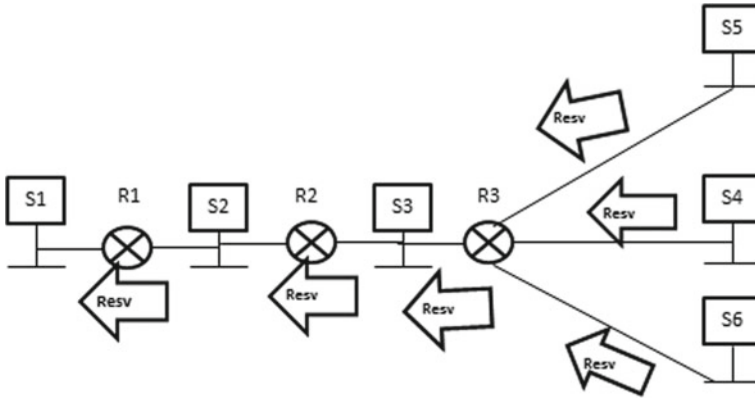
Fig. 4 RSVP RESV messages
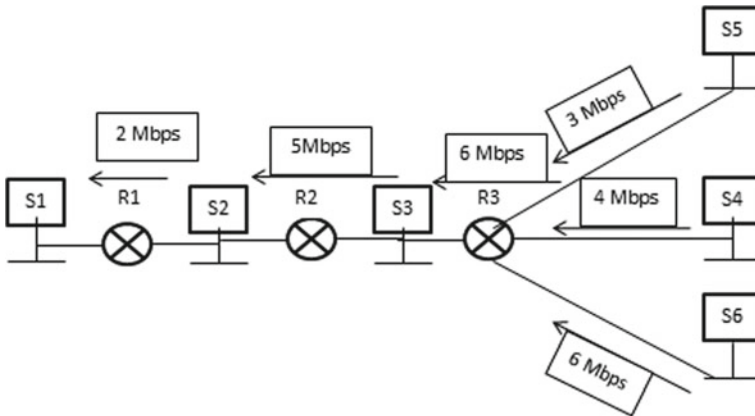
Fig. 5 RSVP bandwidth allocation

The request will be merged, and 2Mbps bandwidth will be given because 2 Mbps can handle both requests, and also stations will not transmit at the same time (Fig. 5).

5 Conclusion

Various channel access methods are studied in this paper for wired and wireless networks. Each method has its merits and demerits. We need to provide priority for delay critical applications. Wireless networks are used very frequently in vast areas of technology. So we need an optimal solution to allocate the bandwidth. More work can be done in the future in this area.

References

1. Mangold S, Choi S, Hiertz GR, Klein O, Walke B (2003) Analysis of IEEE 802.11e for QOS support in wireless LANs. IEEE Wireless Communications. ISBN: 1536-1284/03
2. Choumas K, Korakis T, Tassiulas L (2008) New prioritization schemes for Qos provisioning in 802.11 wireless networks
3. Lagkas TD, Stratogiannis DG, Tsiropoulos GI, Angelidis P (2011) Bandwidth allocation in cooperative wireless networks: buffer load analysis and fairness evaluation. Phys Commun 4:227–236
4. Zuhra FT, Bakar KBA, Ahmed A, Almustafa KM, Saba T, Haseeb K, Islam N (2019) LLTP-QoS: low latency traffic prioritization and QoS-aware routing in wireless body sensor networks. https://doi.org/10.1109/ACCESS.2019.2947337
5. Li B, Huang N, Sun L, Yin S (2017) A new bandwidth allocation strategy considering the network applications. IEEE. ISBN: 978-1-5386-0918-7/17©2017
6. Afzal Z, Shah PA, Awan KM, Rehman ZU (2018) Optimum bandwidth allocation in wireless networks using differential evolution. https://doi.org/10.1007/s12652-018-0858-4

Low Area-High Speed Architecture of Efficient FIR Filter Using Look Ahead Clock Gating Technique

P. Syamala Devi, J. S. Rajesh, A. Likitha, V. Rahul Naik, and K. Pavan Kumar

Abstract Low-area, high-speed and low-power are the primary requirements of any VLSI designs. Finite Impulse Response (FIR) filter is widely uses as an important component for implementing several digital signal processing hardware circuits for their guaranteed linear phase and stability. And consequently, a number of extensive works had been carried out by researchers on the power optimization of the filters. Look Ahead Clock Gating (LACG) is the technique used to reduce the area and delay of the FIR filter circuit. The experimental results show that the proposed FIR filter achieves 74% of delay and area is reduced to 781 from 803 LUT's compared to that using the conventional design.

1 Introduction

Digital finite impulse response (FIR) filter is wide used for enormous unstable signal processing application thanks to its linear-phase property and stability. The downside of the standard FIR filter architectures is that it contains an outsized range of multiplication operations, which needs an outsized computation time and ends up in excessive space and power consumption for period of time applications.

Therefore, analysis on the area-power delay optimized FIR filter design by CSE has gained importance within the recent analysis works. In a hybrid CSE technique has been employed in order to implement an efficient multiplier-less FIR filter. A reconfigurable FIR filter architecture has been proposed in by employing a CSD primarily based VHCSE rule that outperforms the counterpart in terms of reduction in hardware value. Within the case of the SAS, it's essential to implement a low-power, high-speed design of the low pass FIR filter in order that the interval of the SAS are often reduced to supply a warning additional advanced in time whereas intense low power within the period of time sensor-based system. The motivation of this transient lies within the design of area-delay power efficient FIR filter design

P. Syamala Devi (✉) · J. S. Rajesh · A. Likitha · V. Rahul Naik · K. Pavan Kumar
Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: psd@aitsrajampet.ac.in

Fig. 1 FIR filter design
(with n taps)
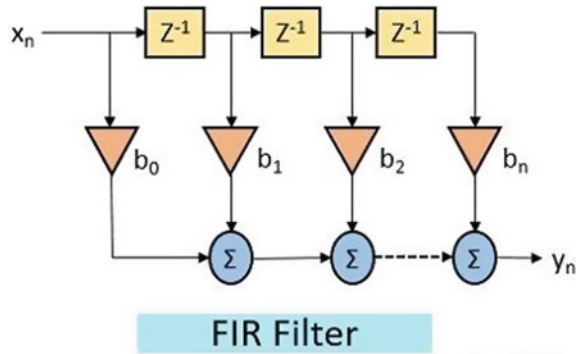
FIR Filter

for preprocessing unstable signal with high exactitude (reduced truncation error) in real-time.

The FIR design is reframed by adders and multipliers utilized in the planning. The planned digital FIR filter style reduces the uses of adders and multipliers blocks for its potency in space and delay. The multiplier interchanged as adders and compressors for its space potency. The input file bits increased by a filter constant so the ripple carry adder selects the actual adder cell to perform addition that depends on carry input. This input bit performs finally of carry save accumulation method, equally next adder cell performs a similar operation. The filtered output constant is obtaining type carry save accumulation method. Thus, the planned ways will minimize the realm, delay of FIR filter style.

An Nth order FIR filter performs N- point linear convolution of the input sequence with filter weights for an input sample.

However, FPGA comes at the price of speed, power and overhead compared to ASICs. The development of the performance of the filter by rule reformulation is proscribed by the generalized reconfigurable nature. For this, many architectures are planned within the last recent years.

A filter is often enforced within the direct form (DF) or the backward direct type (TDF). The backward type and also the direct sort of associate FIR filter are equivalent. It's straightforward to prove that within the direct type, as shown in Fig. 1, the word length of every delay component is adequate the word length of the input. However, within the backword type, every delay component incorporates a longer word length than that within the direct form; what is more, the delay parts are used to delay the product or sum of products. The backward structure reduces the crucial path delay, however it uses additional hardware. Within the crucial path, there are one multiplier $+ (M - 1)$ adders within the DF however only one multiplier $+$ one adder within the TDF. The development on performance is additional noticeable for big M.

Fig. 2 Data driven clock gating
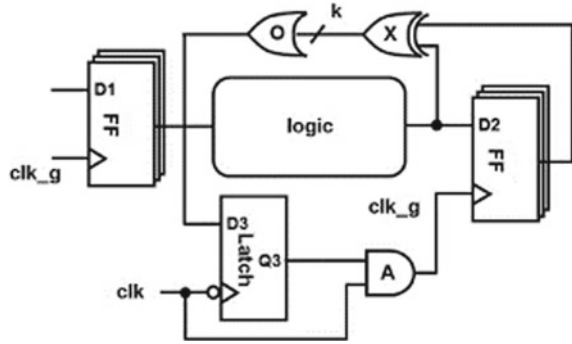
2 Existing Methodology

Clock gating may be a technique that reduces the changing power dissipation of the clock signals. Once this and also the next state of the D flip-flop is determined, it's noticed that once two continuous inputs are identical, the D flip-flop provides a similar worth because the output. Although the inputs don't amendment from one clock to the future, the latch still consumes clock power.

The data driven methods: the clock gating system compares the information values and generates the clock signal (Fig. 2).

Data driven gating has a drawback is it has a very short time window. Whereas the overall delay of the XOR, OR, latch and the AND gate do not exceed the setup time of the FF. The percentage increases with the increase of critical paths in the circuit, suppose if a critical situation arise by downsizing or turning transistors of non-critical path gradually increases to high threshold voltage (HVT) for further power savings. The data driven gating has a complex design methodology when compared to other gating techniques.

3 Proposed Methodology

In proposed system we are using Look Ahead Clock Gating (LACG) technique for providing clock for the FIR filter. And the design of the LACG is as shown in Fig. 3. Look Ahead Clock Gating describes the present cycle data of those FFs on which it depends, the clock enabling signals of each FF one cycle ahead of time. When compared to data driven technique the LACG has a greater advantage of avoiding the tight timing constraints. As like data driven gating, by allocating a full clock cycle are computed for the enabling signals and they are propagated to their gates.
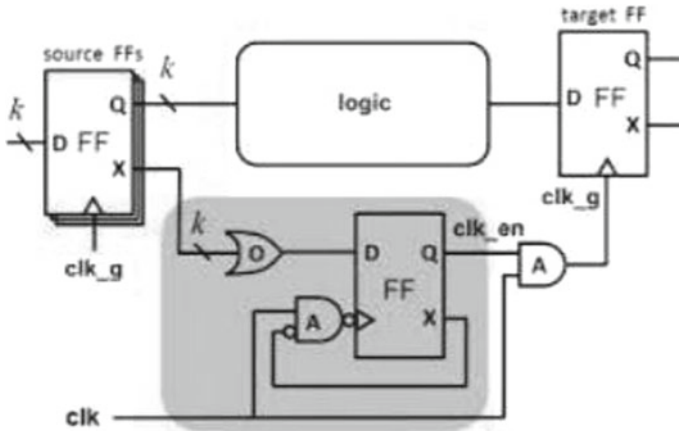
Fig. 3 Practical look ahead clock gating

The data driven gating optimization requires the acknowledgement of flip-flops data toggling vectors in the functional block, whereas LACG is independent of those FFs knowledge. LACG is most effectively used for decreasing the clock switching power and compared to other techniques LACG is preferred mostly. The Look ahead based clock gating is as similar to data driven gating, it reduces the majority of the redundant clock pulses and it reduces the clock switching power. When compared to data driven and auto-gated clock gating, the LACG has a greater advantage of avoiding the tight timing constraints.

4 Results and Analysis

Xilinx ISE is a discontinued software tool from Xilinx for synthesis and analysis of HDL designs, which primarily targets development of embedded firmware for Xilinx FPGA and CPLD integrated circuit (IC) product families. It was succeeded by Xilinx Vivado.

FIR filter will consist of multipliers and adders to provide the output for the input. For example, if we consider the filter design of 4 taps and its weights are 1, 2, 3 and 4. Then the output at every tap for input considered as 10 will be multiplied with the weights and gives as 10, 20, 30 and 40, respectively, for the four taps. Then all the multiplier outputs are given to summer and it provides the result as 100 (Figs. 4 and 5).

Here the results obtained in terms of Area and delay are compared with the previous works and are tabulated in Table 1.
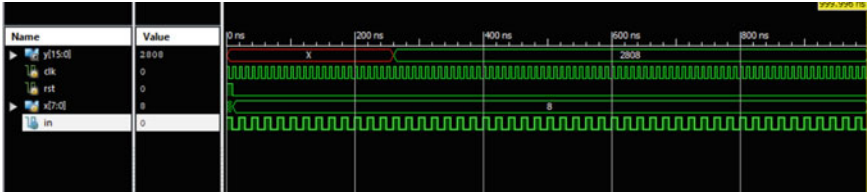
Fig. 4 Simulation results for proposed FIR Filter
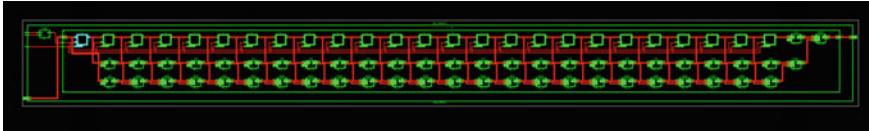
Fig. 5 RTL schematic for proposed methodology

Table 1 Area and delay

Logic utilization	Existing (DDCG)	Proposed (LACG)
Number of Slice and FF LUTs	803	781
Delay	21.85 ns	5.17 ns

5 Conclusion

In this paper, a novel architecture of the FIR filter is proposed. The proposed design is based on the look ahead clock gating (LACG) applied to an appropriate FIR filter structure. To compare the area and delay of the FIR filter we took an eight-bit binary input. In the proposed FIR filter the area and delay are reduced. The proposed filter design reduces the delay up to 45% and the area was reduced to 781 from 803 LUT's.

Bibliography

1. Paliwal P, Sharma JB, Nath V (2019) Comparative study on FFA architectures using different multiplier and adder topologies. Microsyst Technol 1–8
2. Kao H, Hsu C, Huang S (2019) Two-stage multi-bit flip-flop clustering with useful skew for low power. In: Proceedings of the 2019 2nd international conference on communication engineering and technology (ICCET), Nagoya, Japan, Apr 2019
3. Daboul S, Hahnle N, Held S, Schorr U (2018) Provably fast and near optimum gate sizing. IEEE Trans on Comput Aided Des Integr Circuits Syst 37(12):3163–3176
4. Tamil Chindhu S, Shanmugasundaram N (2018) Clcok gating techniques: an overview. In: Proceedings in IEEE conference on emerging devices and smart systems
5. Chandrakar K, Roy S (2017) A SAT-based methodology for effective clock gating for power minimization. Accepted Manuscript in JCSC

6. Hatai I, Chakrabarti I, Banerjee S (2015) An efficient constant multiplier architecture based on vertical-horizontal binary common sub-expression elimination algorithm for reconfigurable FIR filter synthesis. IEEE Trans Circ Syst I Reg Pap 62(4):1071–1080

7. Bhattacharjee P, Majumder A, Das T (2016) A 90 nm leakage control transistor based clock gating for low power flip flop applications. In: 59th International Midwest symposium on circuits and systems (MWSCAS), 16–19 October 2016, Abu Dhabi, UAE, IEEE

8. Singh H, Singh S (2016) A review on clock gating methodologies for power minimization in VLSI circuits. Int J Sci Eng Appl Sci (IJSEAS) 2(1). ISSN: 2395-3470

9. Jensen J, Chang D, Lee E (2011) A model-based design methodology for cyber-physical systems. In: Proceedings of the international conference on wireless communications and mobile computing conference (IWCMC), Istanbul, Turkey

10. Mahesh R, Vinod AP (2010) New reconfigurable architectures for implementing FIR filters with low complexity. IEEE Trans Comput Aided Des Integr Circuits Syst 29(2):275–288

Multimodal Medical Image Fusion Using Minimization Algorithm

Fahimuddin Shaik, M. Deepa, K. Pavan, Y. Harsha Chaitanya, and M. Sai Yogananda Reddy

Abstract By merging anatomical and functional imaging, this multimodal image fusion technique aims to capture certain types of tissues and structures throughout the body. Anatomical imaging techniques can produce high-resolution images of interior organs. If we measure or find medical images independently, such as anatomical and functional imaging, we risk losing essential information. The proposed method or approach, unlike many current medical fusion methods, does not suffer from intensity attenuation or loss of critical information because both anatomical images and functional images combine relevant information from images acquired, resulting in both dark and light images being visible when combined. Colour mapping is conducted on functional and anatomical images, and the resulting images are deconstructed into coupled and independent components calculated using spare representations with identical supports and a Pearson correlation constraint, respectively. The resulting optimization issue is tackled using a different minimization algorithm, and the final fusion phase makes use of the max-absolute-value rule. The image is then normalized, and colour mapping is performed once again by colouring the layers until we obtain the perfect fusion image. This experiment makes use of a number of multimodal inputs, including the MR-CT method's competition when compared to existing approaches such as various medical picture fusion methods. For simulation purposes, the MATLAB R2017b version tool is used in this work.

Keywords Multimodal · MRI · Anatomical · Functional

1 Introduction

The process of joining two or more separate entities to form a new one is known as fusion [1–2]. Medical diagnoses, treatment, and other healthcare applications rely heavily on image fusions. Multimodal medical image fusion is a prominent subfield

F. Shaik (✉) · M. Deepa · K. Pavan · Y. Harsha Chaitanya · M. Sai Yogananda Reddy
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: fahimaits@gmail.com

of image fusion that has made great progress in recent years [3]. Anatomical functional pictures, in particular, have recently been introduced in a study. Anatomical functional images come in a variety of shapes and sizes, each with its own set of features. Image together, in order to express information obtained from multimodal sources images in the same image at the same time to emphasize their respective benefits, in order to carry out complementary information, as well as comprehensive morphology and functional data that reflects physiological and pathological modifications [4].

2   Multimodal Fusion

In order to overcome the limitations of above-mentioned problems, a novel method is proposed on "Multimodal medical image fusion using minimization algorithm" as shown in Fig. 1.

In proposed method, it has two techniques:

1. Functional to anatomical
2. Anatomical to anatomical.

By using this method, we can obtain fused images with a large amount of information.

Using functional and anatomical images as input sources, these two photographs are combined and used to create the colour map [5]. Convert both RGB photographs to greyscale. Using the discrete cosine transformation, decompose the input images [6–8]. Apply the rule of maximal absolute fusion, carry out the picture reconstruction, and apply brightness to an image to boost overall contrast. Perform the standardization process, which increases the greyscale of the given image to a standard
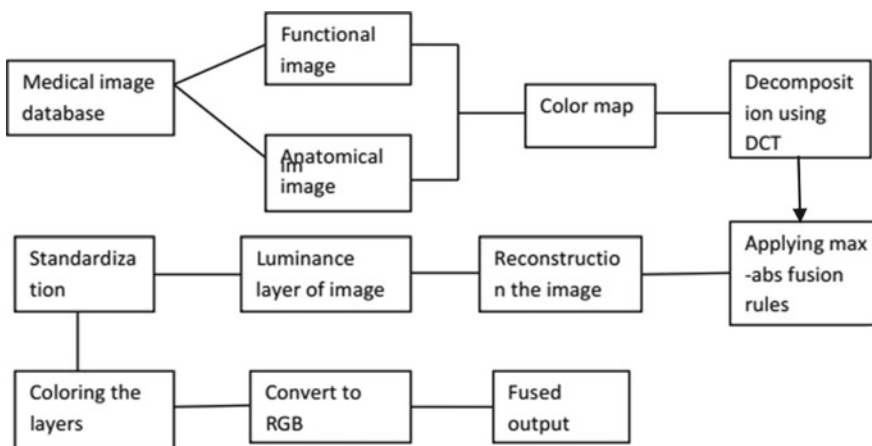
Fig. 1   Block diagram for functional to anatomical

Image 1   Image2   Fused Image

Fig. 2 Functional to anatomical method

Image1   Image2   Fused image

Fig. 3 Anatomical to anatomical method

greyscale [9–12]. It now converts the grey colour to RGB. The final integrated output multimodal medical image is shown.

Consider the two input sources to be anatomical to anatomical images [13]. These two photographs are combined and used to create the colour map. Convert both RGB photographs to greyscale. Using the discrete cosine transformation, decompose the input medical photographs [14–15]. Apply the rule of maximal absolute fusion. Complete the image reconstruction. Perform the standardization process, which increases the greyscale of the given image to a standard greyscale. The final fused output multimodal medical image is shown in Figs. 2 and 3.

3 Results and Analysis

Compared to previous existing methods, the model can generate better fused image without any loss of information. Image fusion has so many contrast advantages; basically, it should enhance the image with all perspectives of image as shown in Table 1.

Table 1 Comparison of various techniques' parameters

Techniques	Average MSE value	Average absolute value	Average gradient	Elapsed time
PCA	10.87	0.009	3.554	–
SWT	5.554	0.003	5.6508	–
DWT	4.4544	0.0038	5.6508	–
HIS	3.3402	–	5.2793	–
Proposed method anatomical to anatomical	$1.040e-04$	$3.105e-03$	–	25.619453
Functional to anatomical	$3.570e-05$	$3.026e-03$	–	8.080941

4 Conclusion

This paper examines many ways to picture fusion. Depending on the application, each method offers advantages and disadvantages. Although these methods improve image clarity to some extent, it has been noticed that the majority of them suffer from colour artefacts and image edge roughness. In the medical imaging industry, more information content and visualization in an image are necessary. In general, wavelet-based schemes outperform classical schemes, particularly in terms of preventing colour distortion, according to the findings of the study. As an image fusion method, SWT outperforms PCA and DWT. To overcome these limitations, we proposed a method of multimodal medical image fusion based on a minimization algorithm. The proposed approach of multimodal medical image fusion, which employs a minimization algorithm, yields higher-quality images with no information loss. The proposed method eliminates distortion in fusion photographs and takes substantially less time to execute than earlier methods. It delivers more information and enhances clarity, helping professionals to quickly examine patients' diagnoses. This approach computes parameters such as average MSE, average absolute value, and elapsed time. The functional to anatomical approach requires substantially less time to perform than the anatomical to anatomical way in this proposed method.

Bibliography

1. James AP, Dasarathy BV (2004) Medical image fusion: a survey of the state of the art. Inf Fusion 19:4–19
2. Li S, Kang X, Fang L, Hu J, Yin H (2017) Pixel-level image fusion: a survey of the state of the art. Inf Fusion 33:100–112
3. Du J, Li W, Lu K, Xiao B (2004) An overview of multi-modal medical image fusion. Neurocomputing 215:3–20
4. Yin M, Liu X, Chen X (2019) Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain. IEEE Trans Instrum Meas

68(1):49–64

5. Du J, Li W, Xiao B, Nawaz Q (2016) Union Laplacian pyramid with multiple features for medical image fusion. Neurocomputing 194:326–339
6. Bhatnagar G, Wu QMJ, Liu Z (2013) Directive constrast based multimodal medical image fusion in NSCT domain. IEEE Trans Multimedia 15(5):1014–1024
7. Jiang Y, Wang M (2014) Image fusion with morphological component analysis. Inf Fusion 18:107–118
8. Du J, Li W, Xiao B (2017) Anatomical-functional image fusion by information of interest in local Laplacian filtering domain. IEEE Trans Image Process 26(12):58555865
9. Yang B, Li S (2012) Pixel-level image fusion with simultaneous orthogonal matching pursuits. Inf Fusion 13:10–19
10. Yu N, Qiu T, Bi F, Wang A (2011) Image features extraction and fusion based on joint spase representation. IEEE J Sel Topics Signal Process 5(5):1074–1082
11. Nabil A, Nossair Z, El-Hennawy A (2013) Filterbank-Enhanced IHS transform method for satellite image fusion. In: IEEE Apr 16–18, 2013, National Telecommunication Institute, Egypt
12. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. SpringerPlus 5:507. https://doi.org/10.1186/s40064-016-2152-2
13. Mirajkar Pradnya P, Ruikar SD (2013) Image fusion based on stationary wavelet transform. Published by International Journal of Advanced Engineering Research and Studies E-ISSN2249–8974
14. Kaur A, Sharma R (2016) Medical image fusion with stationary wavelet transform and genetic algorithm. In: Published by international journal of computer applications (0975–8887) international conference on advances in emerging technology (ICAET 2016)
15. VNMahesh ALLAM, CH NAGARAJU (2014) Blind Extraction Scheme for Digital Image using Spread Spectrum Sequence. Int J Emerg Res Manage Technol 3(28). ISSN: 5681–5688

Pre-collision Assist with Pedestrian Detection Using AI

S. Samson Raj, N. Rakshitha, K. M. Prokshith, and Shumaila Tazeen

Abstract This paper presents a project on a driver assist system which uses the concept of computer vision and a deep learning to detect and identify any pedestrians who suddenly step into oncoming traffic so as to provide an alert or warning to the driver of the vehicle in case of a collision. Since pedestrian-related accidents are increasing every year, with the development of artificial intelligence, machine learning, and deep learning-based pedestrian detection methods has greatly improved the accuracy of pedestrian detection and object tracking becomes much easier. In this project, a deep learning algorithm known as the YOLO algorithm which uses a joint training algorithm is used which helps in the detection and classification of objects. Therefore, with the use of both cameras and the concept of artificial intelligence an efficient pedestrian scanning and identification method can be developed.

Keywords YOLO · Computer vision · Pedestrian detection · Object detection

1 Introduction

Accidents involving pedestrians is one of the leading causes of injury and death around the world. Intelligent driver support systems are introduced to minimize accidents and to save many lives as possible. These systems would first detect the pedestrian and would predict the possibility of collision, and then alert the driver or engage automatic braking or use any other safety mechanism. Advanced driver assistance are electronic systems that aid a vehicle driver while driving. Since, most road accidents occur due to human error these driver assistance systems are developed to automate, adapt, and enhance vehicle systems for safety and better driving. The automated system which is provided by ADAS to the vehicle is proven to reduce fatalities, by minimizing the human error.

S. Samson Raj (✉) · N. Rakshitha · K. M. Prokshith · S. Tazeen
AMC Engineering College, Electronics and Communication Engineering, Bangalore, India
e-mail: sambeckham4@gmail.com

The main objective behind the design of safety features are to avoid collisions or accidents by offering technology based solutions that alert drivers to any potential problems or avoid collisions by implementing certain safeguards and also take control of the vehicle in certain situations. ADAS usually relies on inputs from single or multiple data sources which may include computer vision, lidar, image processing, automotive imaging, radar, and also in-car networking. A pre-collision assist systems uses the concept of deep learning to detect and identify pedestrians who suddenly step into oncoming traffic.

2 Literature Survey

Yamamoto et al. [1, 2] presents project which is mainly based on the concept of lidar technology for efficient pedestrian scanning. The method proposed by Sun et al. in [3, 4] uses classification, clustering, and machine learning techniques for effectively detecting pedestrians, including the application of algorithms such as SVM, neural networks, and AdaBoost for the purpose of distinguishing pedestrians from background. Wu et al. in [5] proposed a method to active cruise control, lane departure warning, blind spot monitoring, and pedestrian detection systems based on sensors such as visible light and thermal infrared cameras, RADARs, or LASER scanners. Lin et al. in [2, 1] proposes a combination of two techniques. The goal is to train the system on the bases of gradients, use the decision tree from which candidate list can be generated with similar features. Also, the SVM trained, that reduces the computational cost and generate the appropriate results. The system proposed by Yang et al. [6, 3] aimed at first, modeling the video surveillance scene via mixed gaussian background model and collecting negative samples from the background images; second, extract the positive and negative samples histogram of oriented gradients (HOG) features, using the local preserving projection (LPP) for dimensionality reduction; and finally, detecting the pedestrian from the input image under the framework of AdaBoost. With the advancement in the field of artificial intelligence, deep learning-based pedestrian detection has greatly increased the accuracy of pedestrian detection. Dang and Wang have proposed in [7] that with the use of YOLO algorithm the pedestrian detection accuracy can be increased with an optimized feature extraction since small objects can also be detected accurately. Lahmyed in [4, 6] aimed at the objects which are present at a greater distance can be detected with use of deep learning based algorithms multi-sensor based detection.

3 Implementation

The working of the project can be divided into three parts as shown in the block diagram, first is the camera subsystem, second is the alert subsystem, and third is the display unit.

3.1 Camera Subsystem

The camera is used to capture the video. The YOLO applies a single neural network to the full image. This network then divides the image into regions and predicts the bounding boxes and probabilities for each region.

3.2 Alert Subsystem

The alert subsystem consists of LED's and buzzers which are used to provide a visual and an audio indication once when the object is detected.

3.3 Display Unit

The display unit consists of a monitor screen which displays the detected pedestrian with the bounding box and the confidence score on top of the bounding box (Fig. 1).
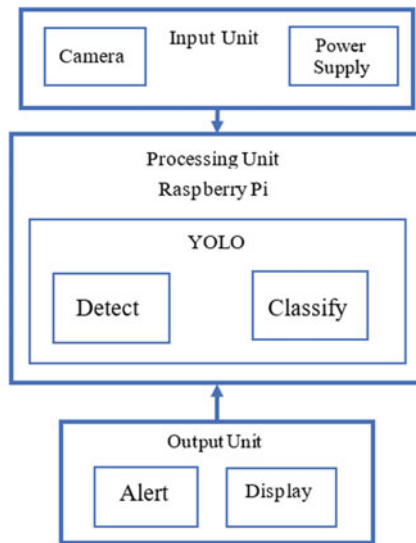
Fig. 1 Block diagram of pre-collision assist system

Fig. 2 Open CV
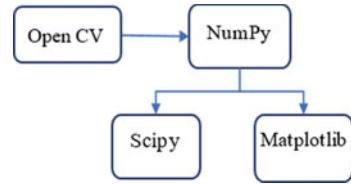
3.4 Working

- Video acquisition–the video is first acquired by the camera which is connected to the raspberry pi.
- Detecting object rectangle by YOLO–Once the object is detected the YOLO draws the bounding box around the detected object with the confidence score displayed on top of the bounding box
- Classifying the object rectangle by YOLO–the detected object is then classified whether it is another car or pedestrian or a cyclist present on the road.
- Tracking the object–once the object is classified its position is monitored by the continuous tracking of the object.
- If the tracked object (pedestrian, cyclist, or dog) enters within the region described in the frame only then the alert is given.
- The alert is given with the help of LEDs and buzzers.

4 Methodology

4.1 Open CV

Computer vision is part or field of AI that deals with how computers can extract or derive meaningful information from various visual inputs such as images, videos, and other sources. Example: Identifying human beings or other objects like lamp post, etc. Computers reads an image in the form of matrix of numbers between 0 and 255.

The Open CV is a library which was designed to solve computer vision problems. In Open CV the array structures are converted to NumPy array which are used by other libraries (Fig. 2).

4.2 YOLO

The working of the YOLO algorithm is shown in the flowchart (Fig. 3).

- As the abbreviation suggests YOLO just looks at the image once.

Fig. 3 YOLO flowchart
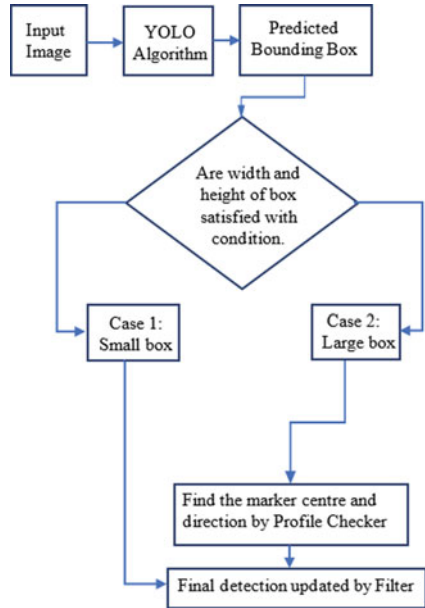
- YOLO first divides an image into a grid which consists of 13 by 13 cells.
- The cells which are present are mainly responsible for predicting a total of 5 bounding boxes. A bounding box is nothing but a rectangular box that encloses and contains an object.
- A confidence score is given by the YOLO algorithm that tells us how certain the predicted bounding box which actually encloses the object. The score does not give any information about the kind or class of object present inside the box but just tells whether the box is good or not.
- The cell also predicts a class for each bounding box, what this does is that it acts like a classifier which means it is able to give a probability distribution of all possible classes.
- To know the type of object present inside the bounding box the class prediction along with the confidence score are combined into one which gives which the final score and also tells us the probability of a specific type of object present inside the bounding box.

5   Results

The result of the project shown in Fig. 4 has no person present within the defined region of the camera frame and the LEDs are not turned on. Figure 5 represents a person within the defined region of the frame and the LEDs and buzzer are turned on thereby providing a visual and an audio alert.

Fig. 4 Image without
pedestrian or person
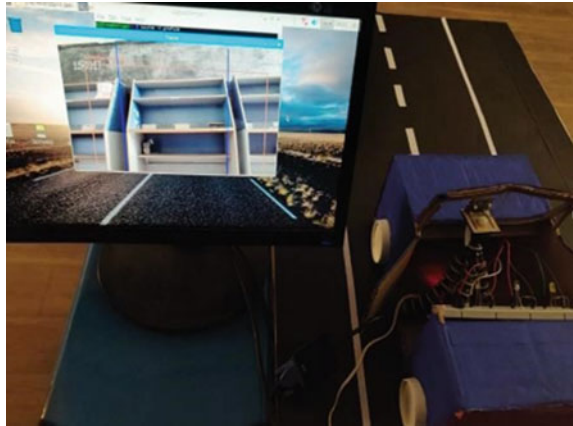
Fig. 5 Image with detected
pedestrian or person
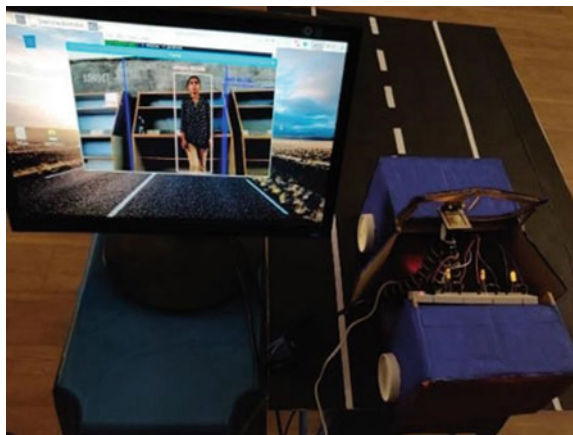
There is also a bounding box present around the detected object or person which also contains a confidence score which tells the percentage confidence of the detected object.

6 Conclusion and Future Scope

This paper deals with the use of the concept of computer vision and deep learning for the detection of pedestrians. It mainly consists of camera which is placed at the dash board for capturing the video input. YOLO algorithm is used to classify the objects detected, YOLO which is installed with Raspberry Pi, Raspberry Pi camera, led, and buzzer is used for detection, classification, tracking, and to send an alert whenever an object is detected. Among the different objects present on the road an

alert is sent only when a pedestrian is detected within a certain region of the frame of the camera. Pedestrians are detected accurately and an early collision alert is sent to the driver of the vehicle.

The application of this project can further be extended for applications such as autonomous driving vehicles, parking assistance systems, traffic sign recognition systems, and lane departure warning systems. The detection accuracy can further be increased with the use of radars or lidars in combination with the camera.

References

1. Yamamoto T, Kawanishi Y, Ichiro Ide I, Murase H (2018) Efficient pedestrian scanning by active scan LIDAR. Nagoya University, Graduate School of Informatics, Aichi Japan
2. Lin C, Lu J, Zhou J (2018) Multi-grained deep feature learning for pedestrian detection
3. Sun W, Zhu S, Ju X, Wang D (2018) Deep learning based pedestrian detection
4. Lahmyed R, Ansari MEL (2016) Multisensors-based pedestrian detection system. In: 2016 IEEE/ACS 13th international conference of computer systems and applications (AICCSA)
5. Wu TE, Tsai CC, Guo JI (2017) LiDAR/camera sensor fusion technology for pedestrian detection. In: Proceedings of APSIPA annual summit and conference
6. Yang Z, Li J, Li H (2018) Real time pedestrian and vehicle detection for autonomous driving. In: 2018 IEEE intelligent vehicles symposium (IV) Changshu, Suzhou, China
7. Lan W, Dang J, Wang Y, Wang S (2018) Pedestrian detection based on YOLO network model. In: Proceedings of IEEE, international conference on mechatronics and automation

An Intrusion Detection System Using Feature Selection Based on Red Kangaroo Mating Algorithm

Soumyadip Paul, Nilesh Pandey, Sukanta Bose, and Partha Ghosh

Abstract Cloud computing takes care of the sudden need of resources and services at any hour of the day from any corner of the world. It is the emerging technology that has made everyday lives much easier. But along with its quickly attained popularity, cloud computing has also gained the attention of attackers. This is why intrusion detection systems are required for protecting the cloud environment. These are installed in host devices or strategic points of a network for detecting attacks and raising alarms to inform the administrative system. In this paper, the authors have proposed an intrusion detection system model using a novel feature selection method. This feature selection method is based on the mating behavior of red kangaroos. When applied on NSL_KDD dataset, the method selects out only 18 features out of 41. The reduced feature set is tested using neural network, decision tree, K-nearest neighbor, random forest, and bagging classifiers. The results prove the proficiency of the proposed feature selection method. Hence, an efficient IDS is developed.

1 Introduction

Cloud computing (CC) is the embodiment of the concept of sharing services and resources according to the requirements of the users. Due to characteristics like high computational power, on-demand service, resource pooling, flexibility, low costing

S. Paul (✉) · N. Pandey · S. Bose · P. Ghosh
Netaji Subhash Engineering College, Kolkata, India
e-mail: sdpaul2000@gmail.com

N. Pandey
e-mail: nileshkk9@gmail.com

S. Bose
e-mail: sukanta.bose@nsec.ac.in

P. Ghosh
e-mail: partha1812@gmail.com

etc., CC has gained tremendous popularity in a very short period of time. But with this growing usage, it has also become a prime target of intruders. So, intrusion detection system (IDS) is developed for defending the cloud from intrusions by taking necessary measures. Based on the source of information, IDSs can be categorized into two types. The first one is host-based IDS (HIDS) which is located at the host machines. This type of IDSs analyzes data obtained from the particular host it is installed at. The other one is called network-based IDS (NIDS) which is deployed at the crucial junction of a network and inspects the inbound and outbound data packets. Again, IDSs can be generally of two types depending on the approach made to identify intrusions. One is misuse detection, and the other is anomaly detection. In case of misuse detection or signature-based detection, the previously recorded attack patterns are considered as signatures. A transmitted data packet is compared with the signatures, and if similarities are found, the data packet is marked as an intrusion. Anomaly detection is applied for finding out new types of attacks. Whenever ample difference is noticed between normal behavioral standards and a data packet, it is pointed as a new type of attack. Numerous IDSs have been constructed till now, but there are some complexities regarding accuracy, false alarms, time consumption, etc. These problems are caused mainly due to the huge size of data, an IDS has to deal with. For this reason, data reduction is necessary. Feature selection (FS) is such a widely used data reduction method. In this paper, the authors have tried to build an efficient IDS using a novel FS method based on red kangaroo mating algorithm (RKMA). The proposed RKMA is a meta-heuristic algorithm that starts with random datapoints in the search space and gradually reaches a solution which is optimum or very close to the optimal solution.

2   Related Works

Cloud has come up as a prominent paradigm to deliver data and services over the Internet. Alongside of its huge utility, cloud has also attained the focus of hackers and attackers. So, for making cloud environment safe and secure, researchers have been building IDSs. Some of these works are described in this section. Machine learning and data mining methods were taken up by Karan Bajaj and Amit Arora for the construction of an IDS in 2013 [1]. Information gain (IG) and gain ratio (GR) as well as correlation between features were used as the basis for FS in their model. They used various discriminative machine learning algorithms for the evaluation of that IDS model.

In the same year, a network anomaly classification technique was proposed by Eduardo de la Hoz et al. [2]. That model selected features having higher discriminative power using fisher discriminant ratio (FDR) as well as performed dimensionality reduction using principal component analysis (PCA), kernel PCA, and isometric mapping. The reduced dataset was then tested using support vector machine (SVM) ensemble.

A NIDS that functions by obtaining optimal feature subsets through a local search algorithm was developed by Seung-Ho Kang and Kuinam J. Kim in 2016 [3]. Firstly, a feature subset was produced by applying k-means clustering algorithm on the training dataset. Then, that reduced feature set was used as a cost function that helped in enhancing the performance of that model.

Again in 2016, Opeyemi Osanaiye et al. worked on an IDS that would efficiently detect distributed denial of service (DDoS) attacks using ensemble-based multi-filter feature selection (EMFFS) technique [4]. In that model, different feature subsets generated with information gain (IG), gain ratio (GR), chi-square as well as ReliefF were combined together by a majority voting system, and decision tree (DT) was applied for classification. The use of nature inspired algorithms is quite common for feature selection purpose in IDS technology.

Seyed Mojtaba Hosseini Bamakan et al. developed an IDS using a modified version of particle swarm optimization (PSO) algorithm [5]. They incorporated chaotic concept in PSO as well as induced time varying inertia weight and time varying acceleration coefficient. That time varying chaos particle swarm optimization (TVCPSO) method continuously adjusted the parameters for multiple criteria linear programming (MCLP) and SVM which executed selection of features. After gaining inspiration from the models discussed above which used various optimization methods, the authors of this paper have focused on developing a novel meta-heuristic algorithm. The proposed algorithm is to be used on intrusion detection datasets for generating optimal feature subsets. These will improve training and testing of IDSs and thus keep cloud intact from attacks.

3 Proposed Model

The authors of this paper have proposed an IDS using a novel FS method. Before deploying an IDS in cloud environment for detecting intrusions, it needs to be trained using datasets. As all features present in a dataset are not equally useful, data reduction techniques can be applied to keep only the important data. In this proposed model, red kangaroo mating algorithm (RKMA) has been applied on the NSL_KDD dataset for FS. Each kangaroo is a datapoint, and each dimension of a kangaroo represents a feature of the dataset. The accuracy of a feature subset is calculated by the fitness value of an individual kangaroo. After completion of the algorithm, the optimal feature subset is obtained from the kangaroo with highest fitness. The flow of the complete proposed model is depicted in Fig. 1.

3.1 Mating Behavior of Red Kangaroo

Red kangaroos, *Macropus Rufus* (also known as *Megaleia rufa*), are social animals living in groups called mobs and are commonly found in the open plains of central

Fig. 1 Flowchart of the
proposed model
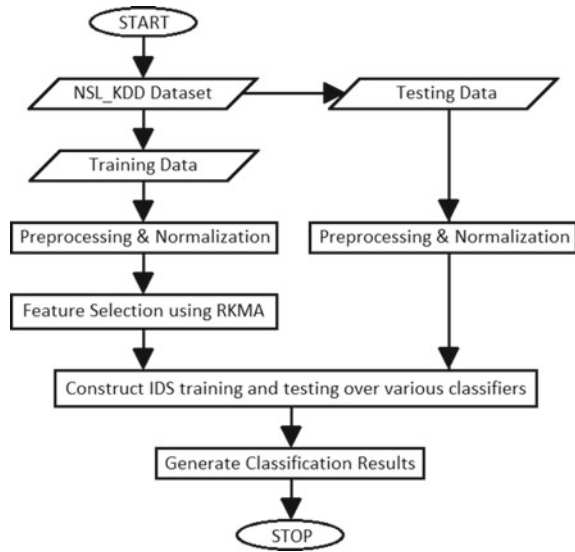
Australia. The males are rusty red in color, and the females show bluish gray coat
[6]. Red kangaroos are polygynous, i.e., one male mates with several females. The
strongest kangaroo of a mob is called the alpha male which leads the group and can
mate with any female. The other males follow the females by sniffing the intensity
of estrone in their urine and look for opportunities to copulate. Boxing takes place if
such two or more males confront while following the same female [7]. The winner of
the boxing gains right over the female. The authors of this paper studied the mating
behavior of red kangaroos and shaped it into an optimization algorithm.

3.2 Red Kangaroo Mating Algorithm (RKMA)

A population with k datapoints is initiated at first. Here, each datapoint denotes a red
kangaroo. Certain equations have been formulated to represent the behaviors of red
kangaroos. For better functioning of the algorithm, some considerations were made
while developing the equations. These are as follows:

- The number of kangaroos in each iteration is kept constant (k). At the beginning
  of each iteration, all kangaroos are considered as unisex. Based on the fitness
  values, they are divided into males (k_m) and females (k_f).
- The alpha male (k_{am}) mates with a selected number of females in each iteration.
  These females do not mate with other males in that iteration.

A random value r_p, lying in the range of [0.3, 0.5], is used to determine the
number of males (k_m). The k_m number of kangaroos with highest fitness values in
the population are considered as males, and the rest are females (k_f). The formula

for generating the male and female population in each iteration is given in Eq. (1) and (2).

$$k_m = \lfloor k \times r_p \rfloor \tag{1}$$

$$k_f = k - k_m \tag{2}$$

In each iteration, the alpha male (k_{am}) copulates with some females at first. These females cannot participate in mating with other males in that particular iteration. Each remaining female kangaroo chooses a number of males determined by a randomly generated integer r_s which depends upon the size of the population. Only the strongest one among the choices of a female gets to mate with her. After copulation, the position of a newborn kangaroo or joey is calculated using Eq. (3).

$$x_{new} = x_f + \left[c_1 \times A_{mf} \times (x_f - x_m)\right] + [x_m \times r_x \times \phi] \tag{3}$$

where x_{new}, x_f and x_m are the positions of the joey, female, and male, respectively. c_1 is a regulation constant, r_x is a random value in the range $[-1, 1]$ used for better exploration, and ϕ is a variable whose value decreases after every iteration. A_{mf} is the attractiveness of a female felt by male calculated by Eq. (4).

$$A_{mf} = I \times e^{-(n \times d_{mf})/c_2} \tag{4}$$

where I is the intensity of estrone in the urine of the female, n is the number of features represented by the datapoint of the female, and d_{mf} is the distance between the two datapoints. c_2 is a constant that regulates the value of A_{mf}.

The joeys are stored in a different set, and after every iteration, they are transferred to the original population. Before starting a new iteration, the fittest k number of kangaroos are considered, and the others are discarded. The division between males and females happens based on Eq. (1) and (2).

As the proposed FS method is applied on a dataset, it is necessary to convert the real values representing the datapoints into binary values. To do so, a hyperbolic tangent function is used which shown in Eq. (5).

$$x_{ij} = \begin{cases} 1 & \text{if rand} < S(x_{ij}) \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where x_{ij} is initially any value denoting the jth dimension of ith datapoint and *rand* is a random value between -1 and 1. $S(x_{ij})$ is calculated using Eq. (6).

$$S(x_{ij}) = \tanh(|x_{ij}|) = \frac{e^{(2 \times |x_{id}|)} - 1}{e^{(2 \times |x_{id}|)} + 1} \tag{6}$$

After completion of all iterations, the kangaroo with the best fitness is taken as the optimal feature subset. The algorithm of the proposed RKMA-based FS method is.

Algorithm. Red Kangaroo Mating Algorithm (RKMA)

Input: NSL_KDD dataset
Output: Optimal feature subset
begin
Generate population with k number of datapoints
Calculate fitness of each datapoint and sort them in descending order
for (iteration < max no. of iterations):
    Separate males and females using Eq. (1) and (2)
    Choose a number of females to mate with the alpha male
    for each chosen female:
        Generate positions of the newborns using Eq. (3) and add in temp dataset
    end for
    for each remaining female:
        Choose a number of males
        Select the male with highest fitness for mating
        Generate position of the newborn using Eq. (3) and add in temp dataset
    end for
    Convert the real valued dimensions of newborns into binary using Eq. (5)
    Combine the newborns with the existing kangaroos
    Sort datapoints in descending order according to fitness values
    Select top k kangaroos for next iteration and clear temp dataset
 end for
end

---

4 Experimental Results

After pre-processing and normalization, the proposed FS method has been applied on the NSL_KDD dataset to evaluate the performance of the IDS model. On applying the RKMA, only 18 out of a total of 41 features are selected. These features are 1, 3, 4, 5, 15, 16, 18, 19, 22, 25, 27, 30, 35, 36, 37, 38, 39, 40. After that, the results have been prepared in two ways. Firstly, neural network (NN), decision tree (DT), K-nearest neighbor (KNN), random forest (RF), and bagging classifiers have been trained and tested using all the records of NSL_KDD train and test dataset, respectively. In the second case, 10-fold cross-validation has been done using the same classifiers as in train-test. Out of them, RF showed the best performance by obtaining accuracy, DR, and FPR of 99.708%, 99.572%, and 0.174%, respectively. The obtained results are as shown in Tables 1, 2 and 3.

Table 1 Performance metrics of train–test using different classifiers

Classifier	Without feature selection			With feature selection		
	Accuracy (%)	DR (%)	FPR (%)	Accuracy (%)	DR (%)	FRR (%)
NN	77.027	61.467	2.410	81.006	69.228	3.429
DT	78.708	68.410	7.682	82.532	73.397	5.396
KNN	77.608	62.417	2.317	78.806	67.186	5.839
RF	77.480	62.620	2.883	78.101	66.282	6.282
Bagging	80.159	70.817	7.497	80.762	69.165	3.913

Table 2 Comparison between train–test results of different IDS models using FS methods

Algorithm	No. of features	Classification accuracy (%)	
		NN	DT
Information gain [4]	13	76.663	79.724
Gain ratio [4]	13	77.089	80.092
ReliefF [4]	13	80.487	81.175
EMFFS [4]	14	77.009	80.820
DMFSM [1]	33	73.550	81.938
TVCPSO-SVM [5]	17	80.935	81.095
TVCPSO-MCLP [5]	17	75.608	77.564
FSNIDS [3]	25	79.183	79.924
Proposed IDS	18	81.006	82.532

Table 3 Comparison between cross-validation results of different IDS models using FS methods

Authors	Classification accuracy (%)	
	Detection rate (%)	False alarm rate (%)
De la Hoz et al. [2]	93.40	14
Kang et al. [3]	99.10	1.2
Bamakan et al. [5]	97.03	0.87
Singh et al. [8]	97.67	1.74
Tavallaee et al. [9]	80.67	NA
Raman et al. [10]	97.14	0.83
Abd-Eldayem [11]	99.03	1.0
Gogoi et al. [12]	98.88	1.12
Proposed IDS (RF)	99.572	0.174

5   Conclusions

From the results, it is clear that the proposed FS method is successful in serving its purpose. In both train–test and cross-validation, the RKMA-based IDS has done better than the other mentioned IDS models. The proposed IDS when deployed in cloud environment will efficiently detect intrusions and make it more secure for users. Thus, the validity of this proposed model is proven.

References

1. Bajaj K, Arora A (2013) Improving the intrusion detection using discriminative machine learning approach and improve the time complexity by data mining feature selection methods. Int J Comput Appl 76(1):5–11. https://doi.org/10.5120/13209-0587
2. de la Hoz E, Ortiz A, Ortega J, de la Hoz E (2013) Network anomaly classification by support vector classifiers ensemble and non-linear projection techniques. In: International conference on hybrid artificial intelligence systems. Springer, pp 103–111.https://doi.org/10.1007/978-3-642-40846-5_11
3. Kang SH, Kim KJ (2016) A feature selection approach to find optimal feature subsets for the network intrusion detection system. Cluster Comput 19:325–333. https://doi.org/10.1007/s10586-015-0527-8
4. Osanaiye O, Cai H, Choo KKR, Dehghantanha A, Xu Z, Dlodlo M (2016) Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing. EURASIP J Wirel Commun Netw 130.https://doi.org/10.1186/s13638-016-0623-3
5. Bamakan SMH, Wang H, Yingjie T, Shi Y (2016) An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization. Neurocomputing 199:90–102. https://doi.org/10.1016/j.neucom.2016.03.031
6. Sharman GB, Pilton PE (1964) The life history and reproduction of the red Kangaroo (Megaleia rufa). Proc Zool Socf London 142(1):29–48. https://doi.org/10.1111/j.1469-7998.1964.tb05152.x
7. Croft DB, Snaith F (1991) Boxing in red kangaroos, macropus rufus: aggression or play? Int J Comp Psychol 4(3):221–236
8. Singh R, Kumar H, Singla RK (2015) An intrusion detection system using network traffic profiling and online sequential extreme learning machine. Expert Syst Appl 42(22):8609–8624. https://doi.org/10.1016/j.eswa.2015.07.015
9. Tavallaee M, Bagheri E, Lu W, Ghorbani AA (2009) A detailed analysis of the KDD CUP 99 data set. In: IEEE, symposium on computational intelligence in security and defense applications, pp 1–6.https://doi.org/10.1109/CISDA.2009.5356528
10. Raman MRG, Somu N, Kannan K, Liscano R, Sriram VSS (2017) An efficient intrusion detection system based on hypergraph—genetic algorithm for parameter optimization and feature selection in support vector machine. Knowl-Based Syst 134:1–12. https://doi.org/10.1016/j.knosys.2017.07.005
11. Abd-Eldayem MM (2014) A proposed HTTP service based IDS. Egypt Inf J 15(1):13–24.https://doi.org/10.1016/j.eij.2014.01.001
12. Gogoi P, Bhuyan MH, Bhattacharyya DK, Kalita JK (2012) Packet and flow based network intrusion dataset. Commun Comput Inf Sci 306:322–334.https://doi.org/10.1007/978-3-642-32129-0_34

Heart Stroke Prediction Using Machine Learning Models

S. Sangeetha, U. Divyalakshmi, S. Priyadarshini, P. Prakash, and V. Sakthivel

Abstract Healthcare field has a huge amount of data. To deal with those data, many techniques are used. Someone, somewhere in the world, suffers from a stroke. When someone experiences a stroke, quick medical care is critical. Heart stroke is the leading cause of death worldwide. Heart stroke is similar to heart attack which affects the blood vessels of the heart. Different features can be used to predict the heart stroke. In order to predict the heart stroke, an effective heart stroke prediction system (EHSPS) is developed using machine learning algorithms. The datasets used are classified in terms of 12 parameters like hypertension, heart disease, BMI, smoking status, etc. These are the inputs for machine learning algorithms which are used to predict the heart stroke. The project aims to build a machine learning model which predicts the heart stroke.

Keywords Heart stroke · Machine learning · Exploratory data analysis · Data visualization

1 Introduction

Stroke has become the leading cause of disability around the world. Around 70 million stroke survivors are expected to live by 2030, with over 200 million stroke adjusted life-years (Dales) lost per year. The burden of stroke was disproportionately high in higher-financial gain nation, with less- and intermediate-financial gain countries experiencing significant growth in the welfare economy. History, testing, laboratory, electrocardiogram, and imaging data are used to deduce and assign causal, philosophy, or phenotypic classification performance in this chaotic stroke sub-type. Ischemic stroke can be caused by a variety of vascular diseases that result in brain thromboembolism. The cause of a stroke is crucial. As a result, it has an impact

S. Sangeetha (✉) · U. Divyalakshmi · S. Priyadarshini · P. Prakash · V. Sakthivel
Vellore Institute of Technology, Chennai, India
e-mail: Sangeetha.s2021@vitstudent.ac.in

P. Prakash
e-mail: prakash.p@vit.ac.in

on the design of epidemiological research and the availability of necessary information to define the sub-type. Expenditure on health care continues to rise, with the most recent estimates being three trillion dollars a year today. Every year, we spend a wide range of age-related, medical errors, and conditions go undiagnosed or mismanaged or misdirected. Heart attack may be the second leading cause of death worldwide, accounting for a quarter of all stroke cases. Current guidelines for the primary prevention of stroke suggest the use of risk prognosis models to diagnose individuals at high risk for stroke (CVD). With early intervention, it is estimated that half of strokes can be prevented by controlling the mutable risk factors in such individuals. It reduces withdrawal and mortality after stroke because many factors can cause a large variety of complications for the patient. All over the past four decades, various risk scores have been presented to diagnose humans at higher risk for cerebrovascular disease. These formulations use conventional statistical ways such as the Cox ratio risk model. It assumes a linear relationship between risk component and therefore the ratio of stroke. When helpful, they assume that the attributes in their models interact in a linear and combination fashion. Some attributes gain profit or loss importance due to the lack or existence of other variables. Consequently, the risk of stroke of those two observations during a linear model is supported. In the most nonlinear model, the possibility is possible due to the presence or absence of homogeneous variables is determined by variable sets of two different types. The latter is arguably better complexity, interactivity, and nonlinearity of reality.

2   Literature Review

In order to predict the stroke, researchers investigate the influence of numerous risk factors on the onset of stroke. Then, using machine learning methods such as random forest, logistic regression, they utilize the processed data to forecast the probability of a stroke occurring. To predict the presence of stroke sickness, the author proposed a model that used classification approaches such as a decision tree, naive Bayes, and neural networks, as well as the principle component analysis approach to minimize dimensionality [1]. The medical institute provides the stroke dataset. The collection includes patient information, medical history, a gene identification illness database, and indication of stroke disease. The data pre-processing approach first eliminates duplicate records, lacking data, disordered data, and illogical data. Following pre-processing, PCA is utilized to reduce dimensions and discover the variables that are more involved in the prediction of stroke prediction illness. The three classifiers are then utilized to diagnose people with stroke illness. In this study, the authors looked at how machine learning algorithms may be used to predict long-term outcomes in ischemic stroke patients [2]. This was a retrospective research that used a prospective cohort to educate acute ischemic stroke patients. At 3 months, favorable outcomes were defined as an altered score of 0, 1, or 2 on the ranking scale. Their predictability has been assessed by developing three machine learning models. A data processing UCO algorithm was proposed in this study, by integrating

three methods: undersampling, clustering, and oversampling [3]. The technology can deal with inclined data from stroke sufferers. In this study, the authors assess the execution of various machine learning models in predicting heart attacks. New outcomes display that RF is the finest model to guess the heart attack possibility on stroke patients' datasets. The precision is 70.05%, and the accuracy is 70.29. To create a prediction model, the authors used a support vector machine (SVM) technique [4]. Also, the SVM method is developed with multiple decision bounds such as linear, quadratic, and cubic. The findings of this study demonstrate that the linear and quadratic SVM performed well in forecasting cardiac stokes through higher correctness values. This paper provides stroke predicting analysis tools based on a deep learning model applied to a heart disease dataset [5]. This study describes predictive analytical tools used for stroke utilizing a deep learning model applied to a dataset of heart illness. The findings of this study are very much accurate than medical counting methods already in usage to notify heart patients whether they are at risk of having a stroke. Because the suggested machine learning technique is intended to attain a degree of accuracy of 99%, this study effort has used an artificial neural network to achieve the greatest efficiency and accuracy. In contrast, the current system predicts strokes using random forest and the XGBoost algorithms. This study aimed to employ artificial neural networks to get the expected outputs [6]. In this paper, the authors created a stroke prediction structure that identifies strokes using actual biosignals and machine learning approaches. The goal of this study was to use machine learning to study and analyze diagnostic procedure of data. The data was cleaned in compliance with the addition criteria that were expressly developed [5].

3 Methodology

The following section explains various steps involved in the proposed system. The step-by-step process of the system is depicted in Fig. 1. It all starts with data collecting. The obtained data is subsequently pre-processed in order to make it usable. The pre-processed data is now loaded into models such as logistic regression and random forest to predict heart stroke in patients. The model is evaluated using accuracy, precision, and a variety of other measures.

3.1 Data Collection

Dataset is organized in a CSV file format which consists of 5110 observations and 12 attributes like hypertension, heart disease, BMI, smoking status, etc. These are the inputs for machine learning algorithms which are used to predict the heart stroke.
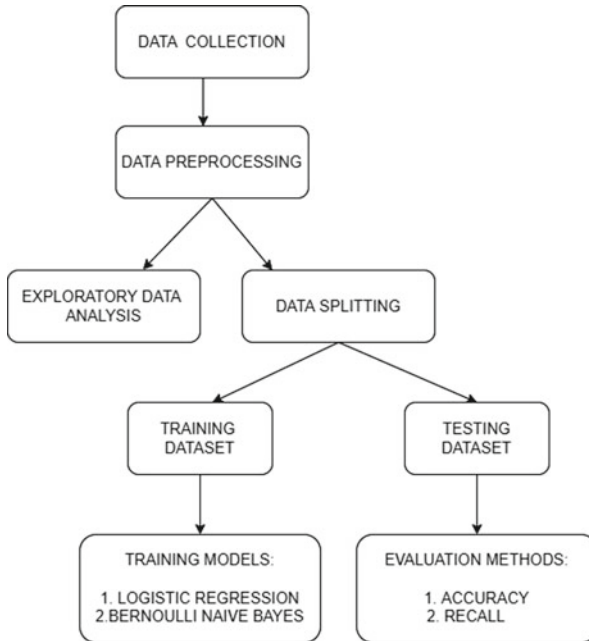
Fig. 1 Flowchart

3.2 Data Pre-processing

Data in its raw form contains incorrect column names; thus, all columns are renamed for better comprehension and convenience of use. As part of the pre-processing, one-hot encoding is employed as the raw data from which categorical characteristics may be derived. All of the characteristics are converted to binary, and the data is transformed to 0 s and 1 s.

3.3 Exploratory Data Analysis

In statistics, exploratory data analysis (EDA) is a method of analyzing datasets in order to sum up their main points, which is commonly done using statistical graphics and other data visualization approaches. Whether a statistical model is utilized, the basic goal of EDA is to examine what the data can tell us without the need of nominal modeling.

Histogram. A histogram is a visual depiction of a numerical or categorical data distribution. A histogram is a type of graph that identifies and displays the fundamental frequency arrangement of continuous data values. This modifies the data to be

tested for its fundamental distribution (e.g., normal distribution), skewness, outliers, and other factors.

Heat map. A heat map is a two-dimensional data visualization tool that displays the ratio of a phenomenon as color (Fig. 2). A heat map is a visual portrayal of data that utilizes code which gives different color to indicate various values. Heat maps are primarily used to improve the amount of outcome inside a dataset and to guide users to the most important sections on data visualizations.

Bar Graph. A bar graph is a chart or graph that shows in rectangular bars using categorical data with heights relative to the values (Fig. 3). A column chart is also known as vertical bar chart.
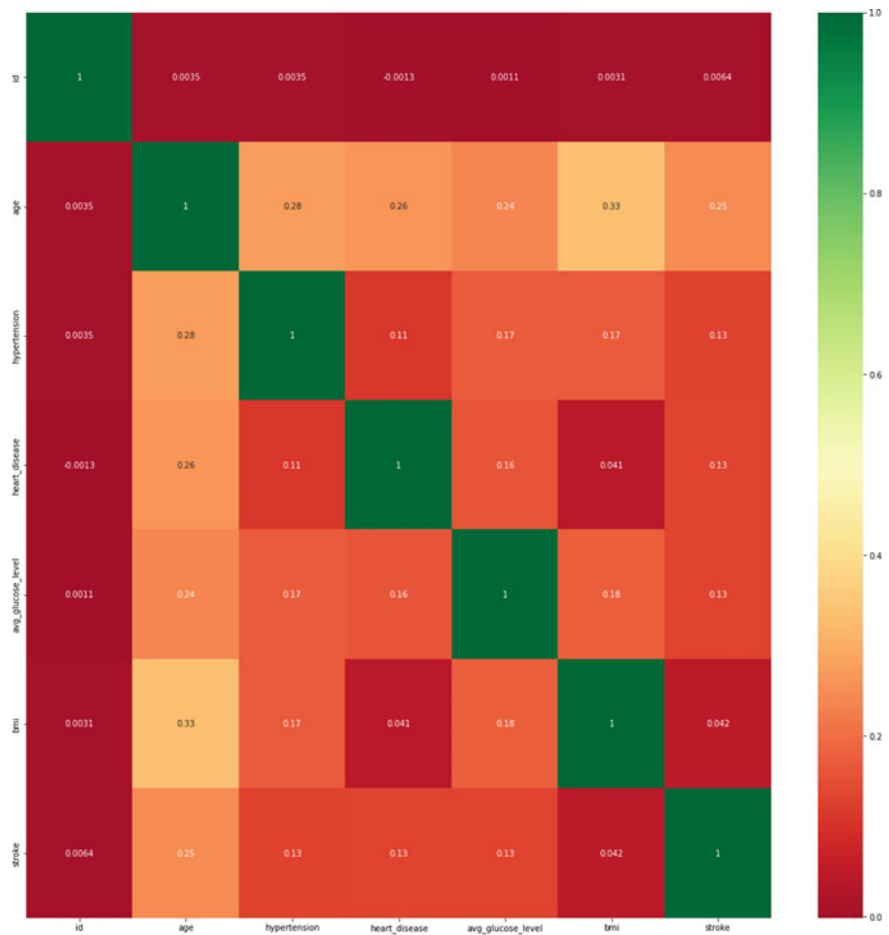
Fig. 2 Heat map

Fig. 3 Bar graph
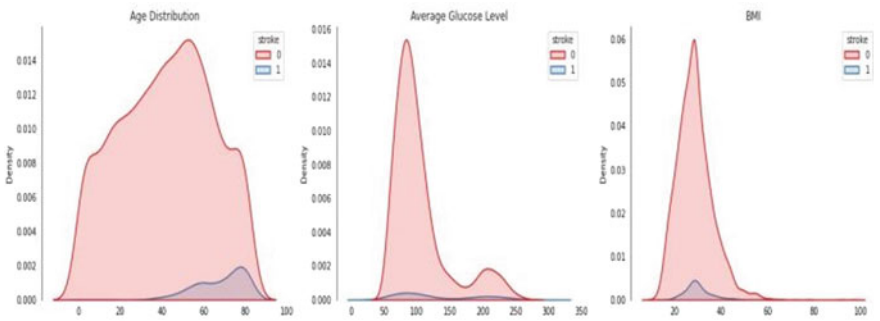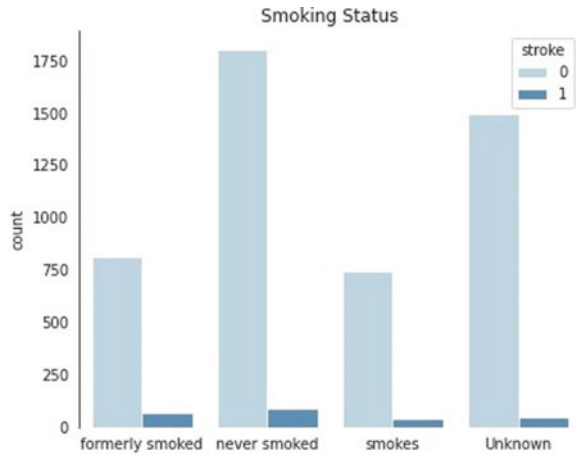
Fig. 4 Density plot

Subplots. Subplots are axes that can be found together in a single matplotlib graphic. The matplotlib library's subplots() function aids in the creation of multiple subplot layouts.

Density Plot. Density plot is nothing but the representation of numerical variable. It shows the probability distribution function of the attributes. Density plot is a smoothened version of the histogram, and it serves the same purpose (Fig. 4).

3.4 Models

Logistic Regression. Logistic regression is used when the dependent variable is dichotomous or binary. Binary logistic regression classification uses one or more predictor variables, which can be continuous or categorical.

Table 1 Comparison table of different models

Classifier	Accuracy
Logistic regression	94.194
Bernoulli NB	92.433

Bernoulli naive Bayes. The Bayes theorem is used to determine the likelihood of any result occurring. Naive Bayes is a machine learning classification approach which is based on the Bayes theorem. The naive Bayes classifier is a probabilistic classifier, meaning it predicts the probability of an input being categorized into all classes given an input. It is also known as conditional probability.

Comparison of Models. Models are compared with train size 0.7 and test size 0.3. Pre-processed data is fed into the model and used for training and testing. It utilizes k-fold cross validation method for cross validation, and the value of the evaluation score shown in Table 1 is the mean score of all the folds. The number of folds is 10.5.

3.5 Metrics for Evaluation

Following metrics are used for evaluation of the models. TP and TN is true positive and true negative values which have same actual and predicted classes, while FN and FP is false negative and false positive which contain different actual and predicted classes. Following equation is used for calculating accuracy for a model.

Accuracy. Accuracy = (TP + TN) / (TP + FP + TN + FN) [7].

4 Results

Table 1 depicts the evaluation of the different classifiers. Logistic regression and Bernoulli navies Bayes give better results compared to other classifiers. Figures 5 and 6 depict the confusion matrix of different classifiers.

In Figs. 5 and 6, it is concluded that logistic regression and Bernoulli navies Bayes have high TP and TN values compared to other classifiers. Figures 7 and 8 depict Receiver Operating Characteristic (ROC) Curve for all classifiers.

5 Conclusion and Future Work

The proposed technology will be used to forecast and detect various illnesses in the future. As a result, it will be valuable in the medical profession. Feature selection can be done in the future when dealing with the prediction. Other machine learning

Fig. 5 Confusion matrix for logistic regression
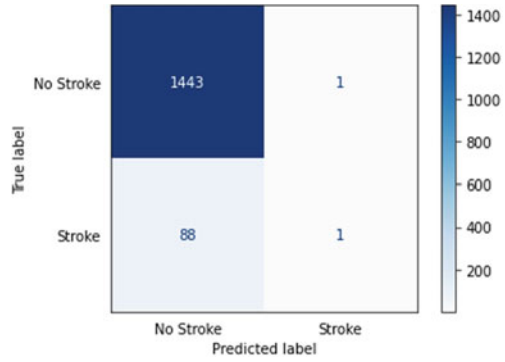
Fig. 6 Confusion matrix for Bernoulli NB
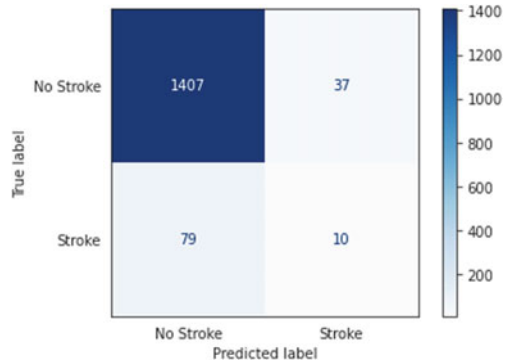
Fig. 7 ROC Curve for logistic regression
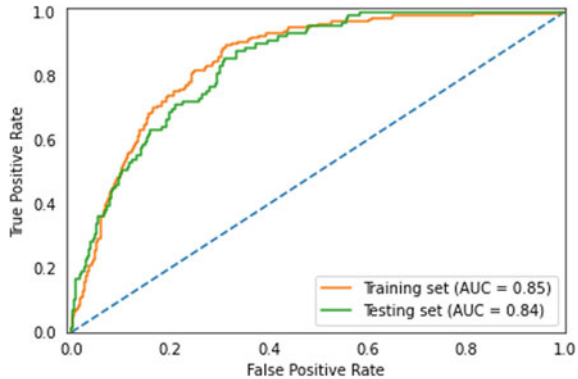
algorithms can also be used for the prediction and to enhance the accuracy. In this paper, two supervised machine learning algorithms are used to predict the heart disease in terms of accuracy. The Bernoulli naive Bayes has predicated the heart stroke with an accuracy of 92%, and the logistic regression model has predicted with an accuracy of 94%. Algorithm which gives the best accuracy is also found.

Fig. 8 ROC Curve for Bernoulli NB
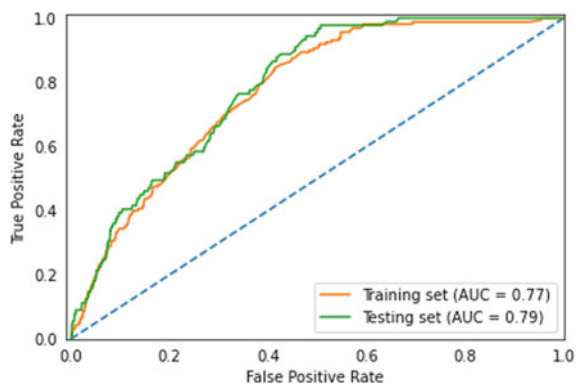
References

1. Sudha A, Gayatri P, Jaisankar N (2012) Effective analysis and predictive model of stroke disease using classification methods
2. Turkmen HI, Karsligil ME (2019) Advanced computing solutions for analysis of laryngeal disorders. Med Biol Eng Compu 57(11):2535–2552
3. Wang M, Yao X, Chen Y (2021) An imbalanced-data processing algorithm for the prediction of heart attack in stroke patients. IEEE Access 9:25394–25404. https://doi.org/10.1109/ACCESS.2021.3057693
4. Puri H, Chaudhary J, Raghavendra KR, Mantri R, Bingi K (2021) Prediction of heart stroke using support vector machine algorithm. In: 2021 8th International conference on smart computing and communications (ICSCC), 2021, pp 21–26. https://doi.org/10.1109/ICSCC51209.2021.9528241
5. Gavhane A, Kokkula G, Pandya I, Devadkar K (2018) Prediction of heart disease using machine learning. In: 2018 Second international conference on electronics, communication and aerospace technology (ICECA), 2018, pp 1275–1278. https://doi.org/10.1109/ICECA.2018.8474922
6. Ponmalar A, Nokudaiyaval G, Vishnu Kirthiga R, Pavithra P, Sri Rakshya RVT (2021) Stroke prediction system using artificial neural network. In: 2021 6th International conference on communication and electronics systems (ICCES), 2021, pp 1898–1902. https://doi.org/10.1109/ICCES51350.2021.9489055
7. Sasikala G, Roja G, Radhika D (2021) Prediction of heart stroke diseases using machine learning technique based electromyographic data. Turk J Comput Math Educ (TURCOMAT) 12(13):4424–4431

Machine Learning-Based Pavement Detection for Visually Impaired People

Swati Shilaskar, Mugdha Dhopade, Janhvi Godle, and Shripad Bhatlawande

Abstract Vision is the most helpful feature of any person in his daily activities or even in his mobility and navigation from one place to the other. But some people born with some visual disabilities find it challenging. One of the major challenges visually impaired people face is to commute through the sidewalks or pavement roads. The proposed work is based on the combination of machine learning and computer vision techniques which can be used to detect the pavement alongside the road. The system will detect the pavement road and give audio feedback to the user. In this work, SIFT feature detector is used. We used five classifiers viz. decision tree, SVM, KNN, logistic regression, and random forest. Among all the models, the highest accuracy was given by the random forest classifier which gave us the accuracy of 83.11%.

Keywords Object detection · Visually impaired · Pavement detection · Sidewalks detection · Object classification

1 Introduction

Visual impairment is a state of a being when a person could not see his surroundings at all or unable to see it partially. Some people are partially blind, and some are completely blind. In short, visually impaired means are those who lose their vision. In this paper, the proposed system is for visually impaired people. Statistically around 2.2 billion people are detected to be visually impaired, from which 80% are aged

S. Shilaskar (✉) · M. Dhopade · J. Godle · S. Bhatlawande
Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune 411037, India
e-mail: swati.shilaskar@vit.edu

M. Dhopade
e-mail: mugdha.dhopade19@vit.edu

J. Godle
e-mail: janhvi.godle20@vit.edu

S. Bhatlawande
e-mail: shripad.bhatlawande@vit.edu

50 years and above and 20% are below 50 years [1]. They face many difficulties in their daily life. Problems like walking around places, studying, getting to know people, etc.

Visually impaired people are mainly at advanced risk for facing a medical problem. Medical care related to blind people received very less consideration as compared to other medical-related problems [2]. Visually impaired people have the similar data requirements as normal people, blind people also need access to important information in their selected reachable format [3]. According to the WHO, the 15% of the world population is struggling with at least one type of disability, such as physical, visual, or hearing impairments [4]. Even when moving from one place to the other, self-path finding is a basic need but for blind persons a solo visit to a new place can not only be very stressful but even unsafe also [5]. Even in the daily activities in their house, they require someone's assistance or a device to track their house surroundings as given in [6]. Nowadays in an advanced automated world, the necessity of self-determining is familiar in case of visually impaired people who are facing many problems. They suffer in unknown environments without any manual aid [7] which could be very troublesome for the people with visual disabilities.

There are many aided devices which help them to tackle their problem. There has been various research and development taking place in this domain to enable the visually impaired to offer a better way to interact with the world. The creation of these electronic and assistive gadgets has led to enhancement of the way of life to these people.

2 Literature Review

Remote controller system [8], wayfinding system [9], recognition system [10]. In paper [8], a smartphone-based multi-device remote controller system is designed. This system uses a smart module which communicates between Bluetooth signals and infrared signals. As Bluetooth signals are omni-directional and IR signals are directional, the person is not required to move a smartphone in a particular way. The advantage of the system is the essence of ICT devices control helps the visually impaired person. The disadvantage of this project is the number of devices controlled by the model are too huge. In paper [9], the independent indoor moves for visually challenged people are concerned. This system uses floorplan and the estimated pose which locate the user's device in a house and also alerts the user by speech command. The advantage of this system is that it takes very short time and gives more accurate results. In the future, the systems implement a loop closure detection method in the previous SLAM system. In paper [10], it presents a 3D object recognition method, and it can be implemented on RNA, i.e., robotic navigation aid, which permits ongoing identification of indoor objects. Inter-plane relationships (IPRs) are extracted by segmenting a fact cloud into numerous planar patches. A Gaussian matrix model is then created to classify each planar patch into one belonging to a particular object model [10]. The method is able to detect structural as well as non-structural objects.

NavCane [11], blind aid stick [12], walking stick [13]. In paper [11], the tool helps blind people to detect a hurdle freeway at outdoor and indoor. It also helps them to recognize an indoor object. Main advantage of this tool is that it offers prior data about any hurdle in the path without producing an overload of information. And it offers low price, and it requires less power embedded devices. In paper [12], the system uses IR sensors, and authors have used ultrasonic range finders for obstacle recognition. A Bluetooth module that is in connection to the GPS technology also to some android app can give audio instructions to some location, and in a difficult situation, it will send SMS to an already defined mobile number. In the future, the system develops better applications using Wi-Fi and IOT so that different features can be predicted. In paper [13], the stick helps visually challenged people to detect the barrier and also provides assistance to reach their end point. The technologies used to design this system are echolocation, image processing, and navigation systems. This uses an ultrasonic sensor to detect the object and echo sound waves. Capturing the images in real time, a sensor that detects the objects is used and a smartphone app is used to navigate the person to the final location using GPS and maps. To design the system mini camera, Bluetooth module, servo motors, microcontroller, and GPS are used.

In paper [14], the system develops a tool which can find different languages like Kannada, Tamil, and English and convert it into audio. The system consists of a Bluetooth headset and a pair of glasses with an HD camera. Image processing is done using the OCR engine tool. The method of voicing alphabets is developed in the device along with appropriate grammar and dictionaries of Kannada, Tamil, and English languages. In paper [15], the project has an indoor map editor and application on Tango devices. Using depth sensor objects are detected. Tango's egomotion tracking can localize orientation on semantic maps, and the barrier finding system creates the safest way to a wanted location. The paper's aim is on dynamic hurdle modeling. In paper [16], the system can detect and identify traffic signs like road crossing, cars, traffic lights, etc. The detection can be based on sensors and stereo camera and output the signal in the audio form. In paper [17], RGB-D camera-based pictorial positioning system for a runtime detection of a robotic location detection aid. By using an RGB-D camera, DVIO judges RNA's pose. From the camera's images, it can extract the floor surface and depth data because of which DVIO has better accuracy.

An electronic travel aid [18] and an electronic travel aid [19] are designed for blind persons. It uses an X86 processor, ultrasonic sensor, and USB camera. The designed system detects the object at the distance of 300 cm, and the USB cam is connected to an eBox 2300™ embedded system which is used for finding the properties of objects, which can also locate a human being. The detection of humans is created on face recognition and texture of clothes. The advantage of the system is that it has less price, portability, and simplicity of controls. In paper [19], the device is designed in the form of eyeglasses which uses depth sensors and ultrasonic sensors to sense a very small obstacle and transparent obstacle in a complicated indoor environment.

Haptic audio and visual techniques provide guided information to visually impaired persons. Advantage of this project is that the device has a low rate and the system is very simple.

Visual assistance [20], blind assistance [21], assistive aid [22], mobility aid [23], visual aid [24]. In paper [20], a system is used to identify the object and sign boards. Raspberry pi, pi camera, and ultrasonic sensor is used to design the system. It can take voice commands and detect the object by using image processing and provides output in the form of audio signal to reach the required object. The system is user friendly that accepts speech signals as input to access his/her basic necessities. In paper [21], the system was designed using a real-time disparity estimation algorithm. Impaired people are using real-time disparity estimation algorithms. The window-based comparing algorithms known as sum of absolute differences (SAD) and zero-mean SAD (ZSAD) are used for the disparity estimation and efficient hardware architectures for those algorithms are implemented in FPGA [21]. From this algorithm, the distance of the nearest obstacle is detected, and the person is alerted via audio signals. The benefit of the system is the system has good optimization and speed. Future, work includes greater improvement in processed image using different procedures or combinations of procedure techniques which best suits the system. In the paper [22], the 360° view camera can take the images from the surroundings and provide the related information to the user via audio signal. This can be done using convolution neural networks (CNN). This system has two challenges: the first is classification and another one is segmentation. The benefit of a 360° camera is that handlers do not need to turn for obstacle detection. The system has gained a 92.8% accuracy. In paper [23], a system is lightweight, portable, and has an ultrasonic sensor and an Arduino to regulate the signals in the occurrence of objects. The sensors and Arduino were not involved in a virtual reality box that was worn by a user which detected obstacles along the path. The detection was done by vibrations from a motor to alert the person of nearby obstacles. In paper [24], the project is designed to detect the obstacles ahead of the person, moving objects, and speed breakers on the ground. The system is also able to detect unexpected falls. The system also informs the present location of the person to keep the information and reports the protector about the person's location. The system has overall accuracy of 98.34%.

Wearable navigation [25], wearable eyewear [26], vision model [1], wearable assistance [27], wearable eyewear [28], intelligent glasses [29]. In paper [25], the system guides users from a location to another location with real-time estimation. A 6-DOF egomotion approximation using spare graphic features are used to design a real-time system. This framework likewise constructs 2D probabilistic inhabitants network maps for proficient navigability examination that is a reason for real-time path planning and barrier saver. The project presented in [26] permits a handler to operate computer-generated 3D objects with a simple. This system used a HWD mounted RGB-D camera. This method divides an object held in hand and judges its 3D location using a camera. The system runs 25 frames per second. This should be possible at 33 ms camera frame rate and 16.6 ms hand detection rate. The disadvantage of this system is that the hand must not block too many image characteristics

since it causes camera tracking to not work. In paper [1], the project uses three ultrasonic sensors to detect the hurdle in every path. The system also detects potholes using ultrasonic sensors and convolution neural network (CNN). The system has 98.73% accuracy at the front side and detects the obstacle at a distance of 50 cm. In this paper [27], the system considers the surroundings with stereo vision and detects the object in the form of audio response. Movable objects are tracked over time. The future extent of the system is to scale down the framework and incorporate the sensors into an eyeglass with the goal that the helmet and the backpack can be eliminated. This will permit visually disabled people to involve it in day-to-day existence. In paper [28], the device can recognize the items from grocery stores, malls, and stores. A button is provided to the device which captures the image with the help of a built-in camera. A CNN is used to train the system. And once the object gets recognized it converts text into speech to inform the user which object, they are holding and also the price of the project. The device is successfully built with the accuracy of 99.35%. In paper [29], the designed system works will be managed by making use of Intel Edison and the framework is designed in Python using OpenCV. Webcam UVC module can capture the images and can be handled by Yocto module which is based on open source in Intel Edison. At the time of image processing, the Yocto module would give related directions from the caught picture. The Bluez module and Alas module can change the guidelines to audio through earphones.

In [30], a development of a complete navigation system that is based on the depth value of Kinect's IR sensor is used which calculates the distance. In [31] that ensures the blind people's interaction with the touch screen devices along with the help of "Braille sketch" by the gesture of the user. In paper [32], the system defines that the visually impaired people find it very difficult to wander in new environments. That is why it requires a system to fuse local 2D grids and camera tracking.

In [33], the proposed model uses a gradient and HLS thresholding for detection of lanes. And a color lane is detected by a sliding window search technique which visualizes the lane. In the future, a real-time system along with hardware execution will be developed. In [34], a vision system for intelligent vehicles is proposed. This system uses the feature of the gray level histogram of the path to detect lane edges. All lane edges are then studied using a decision tree, and finally, the relations between lane edges are studied to create structures defining the lane boundaries. In [35], the system presented can enable lane sensing for different applications. The entire system consists of an image gaining part, a data interface, and a data visualization part, a CMOS camera. The lane detection is based on edge detection method, and Kalman filter is used for parameter approximation.

From the past years, researchers continuously work on different visual aids having different technologies but are having some limitations like in some systems accuracy is less, in some systems the cost of model is more. The weight of the wearables also matters in such cases. Some wearable devices are heavy which can be problematic for users to use on a regular basis. To overcome these limitations and by understanding the difficulty of blind people, we proposed a system which can detect pavement which are fitted on footpaths. This helps them to walk on rush free roads and also avoid accidents.

3   Methodology

This paper presents a mobility aid for detection of pavements. The block diagram of the system is shown in Fig. 1. The system consists of a camera, a processor-based system, and an earphone. The camera acquires information from the surrounding and gives input to processor-based system. The system classifies the images as pavement or non-pavement. The detected output is given via earphone.

3.1   Dataset Collection and Preprocessing

A total of 6043 numbers of images were used (Fig. 2). 40% of images were collected by authors, and 60% were obtained from the Internet. The images of dataset were resized to 180×230 and converted to grayscale.

3.2   Feature Vector Compilation

Scale invariant feature transform (SIFT) was used to extract features from all the images in dataset. SIFT provided feature vector of size $131{,}339{,}216 \times 128$. The process for feature extraction is described in Algorithm 1.
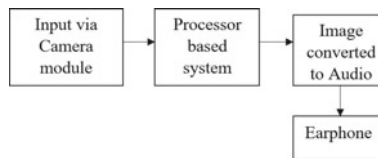
Fig. 1   System block diagram

Fig. 2   Sample pavement images from dataset

Fig. 3 SIFT image descriptor
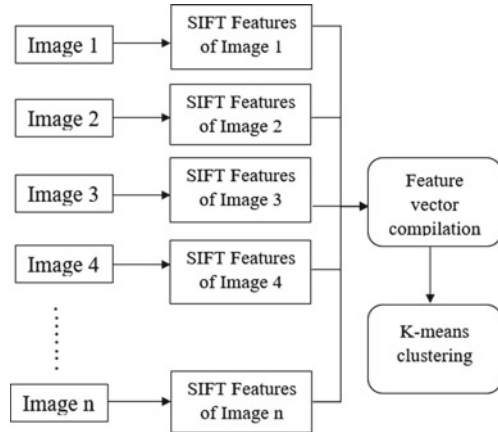
Algorithm 1: Algorithm for SIFT feature extraction

Input: 6043 Images

Output: Feature vector of $131,339,216 \times 128$

1.   Read 6043 images
2.   **for** $i = 0$ to 6042 **do**
3.   Read image
4.   Resize image to 180x230
5.   Convert image to grayscale
6.   Extract SIFT features of image
7.   Features extracted to be stored in csv
8.   **end for**
9.   Obtained feature vector of size $(13,39,216 \times 128)$

The large size feature vector was divided into five cluster ($K = 5$) by using K-means clustering. The value of K was chosen based on elbow method. The SIFT image descriptor has been shown in Fig. 3. The individual SIFT feature vector was predicted using pre-trained K-means model. Obtained histogram was normalized and appended into a csv file to create a feature vector of size 5061×5. This feature vector was further optimized by using principal component analysis (PCA). The PCA provided a feature vector of size 5061×3. The number of principal components was selected based on maximum information content. The process of dimension reduction is shown in Fig. 4.

3.3   Classification and Detection of Pavement

The reduced feature vector was given to classifiers for pavement detection. The optimized feature vector was used to train classifiers such as support vector machine

Fig. 4 Dimensionality reduction
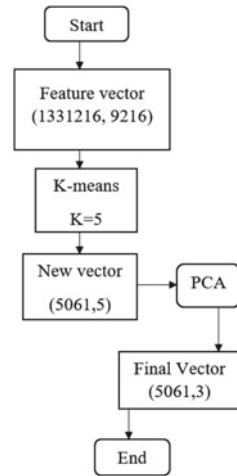
Fig. 5 Classifier used in the system
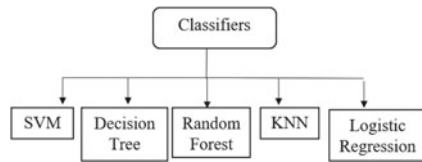
(SVM), decision tree, random forest, KNN, and logistic regression these classifiers are used in the system as shown in Fig. 5.

The first classifier used was SVM. The SVM classifier is a binary classification. The three kernels namely linear, polynomial, and radial basis function (RBF) were used for classification. The hyperplane equation is shown in Eq. (1)

$$f(k) = A(0) + \text{sum}(y_i * (n, n_i)) \tag{1}$$

where k is the input and n_i is the support vector. $A(0)$ and y_i are taken from training data.

Second classifier applied was decision tree. It is a supervised learning classifier based on a tree-like structure with roots and leaves. Parameter used in the model is "entropy". Entropy helps in splitting the data based on the entropy information gain. The entropy is calculated as

$$E(s) = \sum_{i=1}^{c} -p_c c \times \text{Log}_2(p_c) \tag{2}$$

where p_c is probability of class i.

Third classifier applied was K-nearest neighbors (KNN). KNN considers K neighbors to determine the classification of the data point. This model considers the

K points to be five. The distance calculated in KNN is of three types usually, Manhattan distance, Euclidean distance, and finally Minkowski distance. Following is the equation of Minkowski distance. The Minkowski distance is calculated as

$$\text{dist}(a, b) = \left(\sum_{i=n}^{n} |a_i - b_i|^p \right)^{1/p} \tag{3}$$

where $p = 2$.

Fourth classifier used was logistic regression. The algorithm uses probabilities values to classify the data points. The probability values vary between 0 and 1. The output is predicted based on values from 0 or 1 depending on the threshold considered (0.5 in this case). The equation of sigmoid function is

$$f(p) = 1/(1 + e^{-x}) \tag{4}$$

Last classifier was random forest. Random forest is an efficient algorithm that deals with a large number of variable feature selection issues. The advantage of this algorithm is that it effectively handles outliers and noisy data.

The summary of pavement detection is shown in Fig. 6.

The process of classification and detection of pavement is given in Algorithm 2.
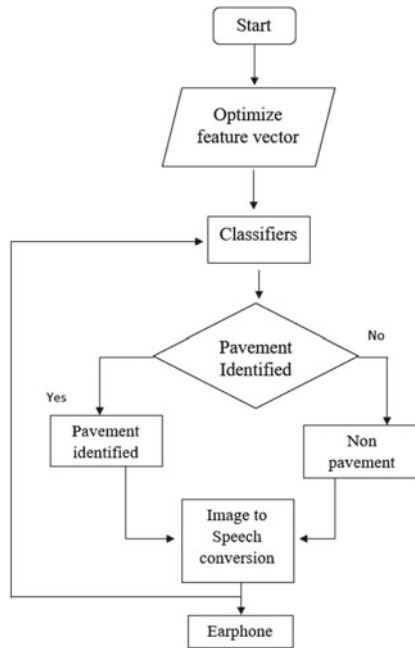
Fig. 6 Flow diagram for classification of pavement

Algorithm 2: Classification and detection of Pavement

Input: Feature vector (5061 × 3)
Output: Pavement detection
1.   Fit the model with training data for classifier
2.   Predict the test data using a pre-trained model.
3.   if (predicted class label == 0 && accuracy >= 0.8)
4.   return (Pavement Road)
5.   else if (predicted class label == 1 && accuracy >= 0.8)
6.   return (Non-Pavement Road)
7.   else
8.   return (Normal Road)

---

4 Results and Discussion

The system was trained and tested with the use of collected dataset. The dataset contained 6043 images. The performance parameter for evaluation used were precision, recall, and $F1$-score. The performance analysis and system accuracy is given in Table 1.

From all the classifiers, accuracy for SVM was 81.34%. The accuracy obtained for random forest was 83.11%. Decision tree gave the accuracy about 77.49%. The accuracy gained for KNN was 78.38%. The accuracy observed for logistic regression was 78.57%. The precision, recall, and F1-score values are presented in Table 1. Each classifier effectively classified the majority of images to their respective classes. As evident from Fig. 7, the precision value and the recall value of random forest provided maximum scores for pavement detection classification.

Table 1 Performance analysis of classifiers

Classifiers	Precision	Recall	$F1$	Accuracy
SVM	85.83	81.96	83.85	81.34
KNN	80.13	82.75	81.41	78.57
Log Reg	80.41	84.30	82.30	77.49
DT	82.48	78.63	80.50	78.38
Rand. Forest	86.89	84.14	85.49	83.11

Classifiers classifiers, *SVM* support vector machine, *KNN* K-nearest neighbor, *Log Reg* logistic regression, *DT* decision tree, *Rand. Forest* random forest
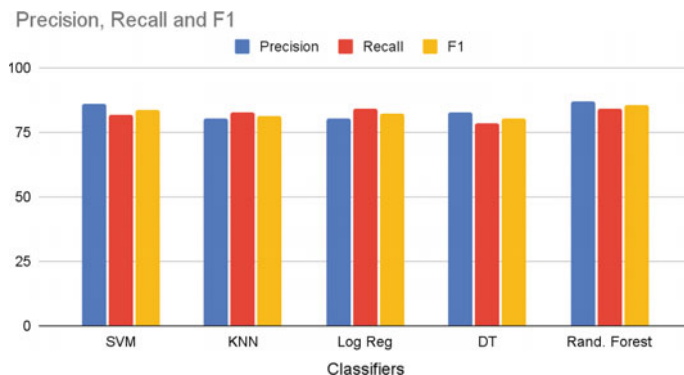
Fig. 7 Performance analysis of classifiers

5   Conclusion

In this paper, we propose a system as an aid to the blind people for detecting pavement road or sidewalk. This will help the visually impaired person to walk alongside the road. Portability of the system makes it easy to carry and offers a hustle free commutation option. With some modifications and addition few more features, the system will be able to detect the curvatures. By training the model further on other objects like trees, uneven pavement, etc., it can give a complete solution for outdoor movement of the visually impaired person. The existing work can be improved by using the above-mentioned features.

References

1. Islam MM, Sadi MS, Bräunl T (2020) Automated walking guide to enhance the mobility of visually impaired people. IEEE Trans Med Rob Bionics 2(3):485–496
2. Zhi-Han L, Hui-Yin Y, Makmor-Bakry M (2017) Medication. handling challenges among visually impaired population. Arch Pharm Pract 8(1):8–14
3. Rayini J (2017) Library and information services to the visually impaired persons. Libr Philos Pract (e-journal) 1510
4. Mancini A, Frontoni E, Zingaretti P (2018) Mechatronic system to help visually impaired users during walking and running. IEEE Trans Intell Transp Syst 19(2):649–660
5. Upadhyay V, Balakrishnan M (2021) Accessibility of healthcare facility for persons with visual disability. In: 2021 IEEE international conference on pervasive computing and communications workshops and other affiliated events (PerCom Workshops). IEEE

6. Chaccour K, Badr G (2016) Computer vision guidance system for indoor navigation of visually impaired people. In: In 2016 IEEE 8th international conference on intelligent systems (IS). IEEE, pp 449–454
7. Felix SM, Kumar S, Veeramuthu A (2018) A smart personal AI assistant for visually impaired people. In: 2018 2nd International conference on trends in electronics and informatics (ICOEI). IEEE
8. Rao SN, Suraj R (2016) Smartphone-aided reconfigurable multi-device controller system for the visually challenged. In: 2016 IEEE international conference on computational intelligence and computing research (ICCIC). IEEE
9. Zhang, He, and Cang Ye (2017) An indoor way finding system based on geometric features aided graph SLAM for the visually impaired. IEEE Trans Neural Syst Rehabil Eng 25(9):1592–1604
10. Ye C, Qian X (2017) 3-D object recognition of a robotic navigation aid for the visually impaired. IEEE Trans Neural Syst Rehabil Eng 26(2):441–450
11. Meshram VV et al (2019) An astute assistive device for mobility and object recognition for visually impaired people. IEEE Trans Human-Machine Syst 49(5):449–460
12. Arora AS, Vishakha G (2017) Blind aid stick: hurdle recognition, simulated perception, android integrated voice based cooperation via GPS along with panic alert system. In: 2017 International conference on nascent technologies in engineering (ICNTE). IEEE
13. Krishnan A et al (2016) Autonomous walking stick for the blind using echolocation and image processing. In: 2016 2nd International conference on contemporary computing and informatics (IC3I). IEEE
14. Ramkishor R, Rajesh L (2017) Artificial vision for blind people using OCR technology. Int J Emerg Trends Technol Comput Sci (IJETTCS) 6(3)
15. Li B et al (2016) ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind. In: European conference on computer vision. Springer, Cham
16. Ye W et al (2010) The implementation of lane detective based on OpenCV. In: 2010 Second WRI global congress on intelligent systems, vol 3. IEEE
17. Söveny B, Kovács G, Kardkovács ZT (2014) Blind guide-A virtual eye for guiding indoor and outdoor movement. In: 2014 5th IEEE conference on cognitive infocommunications (CogInfoCom). IEEE
18. Kumar A et al (2011) An electronic travel aid for navigation of visually impaired persons. In: 2011 Third international conference on communication systems and networks (COMSNETS 2011). IEEE
19. Bai J et al (2017) Smart guiding glasses for visually impaired people in indoor environment. IEEE Trans Consum Electron 63(3):258–266 (2017)
20. Jain BD, Thakur SM, Suresh KV (2018) Visual assistance for blind using image processing. In: 2018 International conference on communication and signal processing (ICCSP). IEEE
21. Sekhar VC et al (2016) Design and implementation of blind assistance system using real time stereo vision algorithms. In: 2016 29th International conference on VLSI design and 2016 15th international conference on embedded systems (VLSID). IEEE
22. Ali M et al (2017) 360° view camera based visual assistive technology for contextual scene information. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC). IEEE
23. Zahir E et al (2017) Implementing and testing an ultrasonic sensor-based mobility aid for a visually impaired person. In: 2017 IEEE Region 10 humanitarian technology conference (R10-HTC). IEEE
24. Rahman MM et al (2020) Obstacle and fall detection to guide the visually impaired people with real time monitoring. SN Comput Sci 1:1–10
25. Lee YH, Medioni G (2016) RGB-D camera based wearable navigation system for the visually impaired. Comput Vis Image Underst 149:3–20
26. Ha T, Feiner S, Woo W (2014) WeARHand: head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception. In: 2014 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE

27. Schwarze T et al (2016) A camera-based mobility aid for visually impaired people. KI-Künstliche Intelligenz 30(1):29–36
28. Pintado D et al (2019) Deep learning based shopping assistant for the visually impaired. In: 2019 IEEE international conference on consumer electronics (ICCE). IEEE
29. Rani KR (2017) An audio aided smart vision system for visually impaired. In: 2017 International conference on nextgen electronic technologies: silicon to software (ICNETS2). IEEE
30. Kanwal N, Bostanci E, Currie K, Clark AF (2015) A navigation system for the visually impaired: a fusion of vision and depth sensor. Appl Bionics Biomech 2015
31. Aslam SM, Samreen S (2020) Gesture recognition algorithm for visually blind touch interaction optimization using crow search method. IEEE Access 8:127560–127568
32. Díaz-Toro AA, Campaña-Bastidas SE, Caicedo-Bravo EF (2021) Vision-Based system for assisting blind people to wander unknown environments in a safe way. J Sens 2021
33. Haque MR et al (2019) A computer vision based lane detection approach. Int J Image Graph Sign Proces 12(3):27
34. Gonzalez JP, Ozguner U (2000) Lane detection using histogram-based segmentation and decision trees. In: ITSC2000. 2000 IEEE intelligent transportation systems. Proceedings (Cat. No. 00TH8493). IEEE
35. Goldbeck J, Huertgen B (1999) Lane detection and tracking by video sensors. In: Proceedings 199 IEEE/IEEJ/JSAI international conference on intelligent transportation systems (Cat. No. 99TH8383). IEEE

A Review on Location-Based Routing Protocols in Wireless Sensor Networks

K. Md. Saifuddin and Geetha D. Devanagavi

Abstract Nowadays, location-based wireless sensor network (WSN) routings are attracting a lot of attention throughout the research field, especially due to their scalability. Sensor networks offer a comprehensive and integrated taxonomy of location-based routing protocols. The whole paper is aimed at contrasting several local routing protocols within networks with wireless sensors. Though localized routing protocols seem to be constrained, multi-hop data transfer may be facilitated. In order to relay position information, location-based routing protocols also require a lot of resources. We define protocols on either the grounds of scalability, energy consciousness, data aggregation, QoS, and versatility.

Keywords Location-based protocol · Scalability · Mobility · Wireless sensor networks

1 Introduction

Location-based routing also recently emerged as the largest field throughout wireless sensor networks (WSNs). There seems to be no IP addresses in sensor nodes; thus, IP-based protocols for sensor networks would not be used. Because of scarce capital and the complex existence of sensor networks, the construction of an efficiently, scalable, and easy protocol is quite difficult. When using the position information, that node doesn't have to confuse calculations to locate the next move in place. Location-based protocols are very effective when routing data packets since they benefit from pure localization knowledge rather than global information about topology.

K. Md. Saifuddin · G. D. Devanagavi (✉)
Reva University, Bangalore, India
e-mail: dgeetha@reva.edu.in

K. Md. Saifuddin
e-mail: saifu426@gmail.com

The survey and taxonomy of location-based sensor network travel are provided in this article. A node with a packet through location-based routing from each data packet sends (Source) and adds another destination (Sink). This packet becomes obtained being sent to the next neighbor through one locally closest to both the destination by intermediate nodes throughout this direction. The method persists until another destination node receives the data packets. Due to all the locality, minimum condition needed for each node to only be sustained is minimum.

That has less overhead communication, since routing tables advertising are not required, as in conventional routing protocols. Location-based routing thus would not need routes to be built or retained. Thus, the location-based routing preserves energy and bandwidth, as a one-hop distance request message and state propagation are indeed not necessary. Place-based routing utilizes nodes to even provide location data, greater scalability, and performance.

The localized routing needs the right position information, and any localization process may be used to acquire that information. For several wireless network applications, position knowledge is important, so it is anticipated that certain localization devices would be fitting from each wireless sensor node of the network. Through position sensing on the basis of proximity or through satellite, acoustic or infrared signals, many techniques are available. These strategies vary in granularity, range, sophistication, and cost of implementation. That flooding becomes used most of the position structures to expand that sink to many other node positions, which for broad WSNs in particular are unwanted where the various mobile sinks but sources occur.

Location-based routing typically utilizes a communication method to send an incoming packet to such a destination through source. Greedy reaches the friend, nearer to either the destination, with forward packets. This implies that perhaps the network becomes dense enough; each node provides its own precise position information, the coordinates of its neighbors, including high reliability. Any WSN implementations may well be acceptable through dense sensor distribution and relatively precise position details. In some practical use, however, the high reliability of the connection is not appropriate. Since recent findings indicate that wireless networks can be very unstable and must deal through protocols at higher standards. There are several advancement techniques designed to boost regional routing efficiency. These transmitting techniques may be usually separated into two categories: expanse and receipt. Whereas nodes have to distinguish their neighbor's distance in either a remote strategy, that nodes often know their competitors' packet delivery rates mostly in reception strategy. Together techniques utilize the greedy transmission pick the ensuing phase throughout that transmission method.

2   Background

In a number of uses, including automotive, civil, and military, the wireless sensors are used. Any of the following are listed:

Military applications: Cameras are commonly used during military applications. Implementations include tracking, connectivity to reference stations from intractable places. Because they are cheap and are used in vast quantities, losses of either of these sensors will not hinder their use.

Distributed surveillance: High mobile sensor networks such as the Zeus, operating Odyssey Marine Exploration Inc. enable vast volumes of low power information to be transmitted.

Structure monitoring: Structural surveillance identifies, locates, and estimates the harm level. Structures throughout civil engineering with the aid of such testing could be checked for fitness.

Pollution and toxic level monitoring: Devices gather data from agricultural environments and areas whereby poisonous leaks take place. This is beneficial for both the sensing and transport to distant stations for study of radioactive, chemical, and biological phenomenon in the area.

Rainfall and flood monitoring: Those networks shall contain water level, wind, and temperature sensors and shall be sent to the central research and weather prediction database.

Smart sensor networks: There are a variety of independent sensors in such networks. Any single sensor, it decides locally and then integrates these decisions and tests mostly on basis of a particular algorithm that gives rise to such a global decision.

Other applications: Ecosystem monitoring with bio-complexity is part of several applications; others provide mining and metal analyses as resources explorations.

Health applications provide patient identification and medicinal administration control in hospitals. Both household appliances and smart home and workplace settings there are fantastic market prospects.

In general, wireless sensor networks can be used for environmental tracking. Throughout this primary field of study, the key challenges are sensor energy conservation and data enhancement. One of the main problems in the field of WSN research has been the selection of the required routing protocol. The consumer wants to evaluate various styles of routing protocols before choosing the routing protocol. Routing protocols will, depending mostly on network configuration, be split in flat hierarchical and localized router protocols. Both sensors are also in flat protocols [1]. The same tasks have been allocated. Various energy level sensors are given various functions in hierarchical network architectures. The data transmission position knowledge of the sensors is used in location-based routing protocols.

Local routing protocols [2] lose resources for transmitting position information's such that data are transmitted and according to route planned. Although it looks more complicated, it can send data through various sensors. If another range of transmission is restricted, local routing protocols remain much additional helpful.

3   Location-Based Routing Protocols

Research paper explores state-of-the-art destination protocols for WSNs [3]. That Internet Protocol (IP) addresses of sensor nodes could not be used, so that the IP protocols with sensor networks could not be used. Based on minimal resources as well as the complex character of the sensor network, it's indeed extremely difficult to create an effective, scalable, and easy WSN protocol. That node does not require complicated calculations to determine the next move during location-based routing, because routing decisions use position details. Place protocols [4] are extremely effective that routing data packet, as they are using pure position data instead of global topology information, utilizes the location between nodes to render the protocol more effective and scalable. Three facts are required. Second, every network node needs to recognize its own GPS or even other tools through location information. Secondly, any node should be informed of the position of its neighboring nodes one-hop apart. Fourth, the position of both the destination node should be identified to the source node.

Routing protocols dependent on position may be predominantly categories as seen in Fig. 1. Many protocols focused on positions are sending the packet to its destination utilizing greedy algorithms. Such algorithms vary except for how they cope with the issue of whole communication. In this section, we survey in depth the protocols which fell just below WSN location-based routing.

A. **Geographic Adaptive Fidelity**

Each protocol is energy-awareness-based. Only nodes not really in usage are disabled to save electricity. This is a three-step protocol. First of all, the direction is found in the shape of even a grid. Secondly, any active node claims it is part of the routing, though no additional energy is needed to locate the path. Finally, their packets are passed to the target. When the radio is shut down as well as the routing is switched from over nodes into sleep mode. Both latency and package failure, GAF accomplishes well with an ordinary adhoc routing protocol and increases the network life through energy saving.

Fig. 1 Categorization of location-based routing protocols
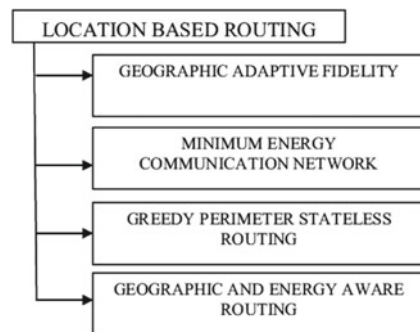
B. **Minimum Energy Communication Network**

MECN was a protocol dependent on position to accomplish minimum random network electricity, minimum energy, using mobile sensors that hold a minimal network of electricity. This method of routing builds a different sub-network for which there is a master node that has capacity for the current non-active nodes using a very lower power topology. There are fewer nodes throughout this sub-network. Such nodes require fewer resources to connect with either the network nodes. There are two steps of MECN routing protocol during packets movement. During the first stage, a locking graph with local measurement throughout the nodes as well as an optimal energy relation is formed. And then, throughout the second level, the relationship is made and the Bellmen Ford Shortest Path Algorithm is used for transmitting minimal energy information.

C. **Greedy Perimeter Stateless Routing**

Greedy Perimeter Stateless Routing (GPSR) [5] favors the continuing usage for a scalable network with geographical locations. The dilemma of gaps that are filled by the planar graph is really in spatial routing. GPSR is really the oldest technique in flat graph routing. That planar diagram perimeter is often used to identify GPSR roads. The nodes operate in multiple states, which have been minimized by the GPSR routing system. It can be used to find and classify the node position facility.

The algorithm takes two forms: selfish and greedy. Greedy shipping is applied throughout the region of which greedy shipment is not willing to participate, but transmission of the periphery. Greedy forwarding becomes achieved in the standardized topology by understanding the new neighbors. If gullible path forwarding isn't really feasible, another route becomes performed using local knowledge on topology. Geographical routing provides all the benefits of GPSR. It uses knowledge from the next node to include a packet. Through GPSR, their states are relative to the number between neighbors, the origins of traffic as well as the DSR to something like the no. of routes experienced as well as the path duration in these specific hops.

The geography is used to obtain a limited state per node routing through the GPSR routing algorithm [5]. GPSR produces protocol routing traffic throughout amounts regardless of network duration. GPSR advantages are all based on regional routes and are used for transmission decisions only through immediate surrounding details.

D. **Geographic and Energy Aware Routing**

GEAR is really a WSN protocol for geographical location [6]. It is indeed a protocol that is energy efficient. That lifetime of GEAR's network is greater than other geographical routings. A specific tool, called Geographic Information System that only functions in some segments of the system, has been used in GEAR for position of both the sensor nodes. That Geographic Information System also allows pick sensor nodes and for packet to be transmitted A process named the conscious heuristic destination. As position hardware, GEAR also uses GPS. GEAR utilizes the collection from energy neighbors for transmission of the message in such a geographical

area and utilizes a recursive geographical transmission or limited algorithm which disseminates that packet to safely transmit the packet.

With both the assistance of the selected nodes which are geographically n energy conscious and are therefore circulating that packet throughout the field, key routing measures are to submit packet to its destination. Next, it uses the chosen regional and energy-conscious nodes to provide further packets to its destination to reach minimal costs. If another neighbor is closer to goal area, the closest hop nodes would be picked and even if the neighbor is absent, a node becomes chosen with a next hop. Second, if the packets will then be disseminated, h(N, R) would have been the cost learned for area R, the N state, and R region of both the goal. Each node proceeds to change its costs by neighbor.

The node tests the expense of almost all of the neighbors in this same area before sending a packet either a region. When a node doesn't really cost a neighbor to such an area to be learned, so the approximate cost is measured below:

$$c(\text{Ni}, R) = ad(\text{Ni}, R) + (l - a)e(\text{Ni})$$

where a = tunable weight, from 0 to 1.

d (Ni, R) = normalized the largest distance among neighbors of N.

$e(\text{Ni})$ = normalized the largest consumed energy among neighbors of N.

Because as node needs that packet to both the endpoint region is examined for something like a neighbor nearest to the destination area and where there is upward of one neighbor, the studied costs are reduced (Nmin, R). Anyway, it sets its own expense:

$$h(N, R) = h(N_{\min}, R) + c(N, N_{\min})$$

c (N, $N_{\min}$) = the transmission cost from N and $N_{\min}$

GEAR is utilizing a recurring transmitting system for transporting packets since basic flooding of electricity is incredibly costly. It offers a scenario about how to use equilibrium resources, which enhances the longevity of both the networks. GEAR not just to decreases route setup energy usage but also in terms of information distribution, it is stronger than GPSR. The communication throughout the network increases after partitioning of nodes. The routing algorithm should really be built to minimize energy throughout data transmission as even more energy than in sensing has been used to transfer.

4   Comparison of Location-Based Protocols

Table 1 demonstrates the thorough contrast between the algorithms alluded to here. Mobility, energy perception, self-reconfiguration, and negotiate are the contrasted properties Data aggregation, service consistency (QoS), sophistication of state, scalability and embedding. In comparison with both the static nodes, routing algorithms

Table 1 Comparison of different routing protocols

Protocols	Mobility	Energy aware	QoS	Data aggregation	Scalability
GAF	Yes	Yes	No	No	Good
MECN	Yes	Yes	No	No	Limited
SMECN	Yes	Yes	No	No	Limited
GPSR	Yes	Yes	No	No	Limited
GEAR	Yes	Yes	No	No	Limited
ALS	No	Yes	No	No	Good
TBF	No	No	No	No	Limited

that involve mobile nodes were dynamic in design. Due to limited resources accessible from the small wireless sensor nodes, energy knowledge in localized routing protocols is indeed an essential feature of WSN. Since these protocols include position details, a localizer such as a GPS is needed to provide wireless sensor nodes or even other techniques [7–9] to provide locations information. For localized routing protocols, scalability and QoS are about as critical as every other resource. Many localized routing protocols were auto-configurable, because even though certain nodes on the network are down, that network will rearrange and run with little or no external assistance (non-functional).

5 Conclusion

This paper contrasted a variety of routing protocols of wireless sensor networks. Multi-hop routing protocols were location-based. While this form of routing protocol has restricted versatility and scalability, the transmission spectrum may be applied if the transmission becomes limited. In other terms, it is really important to study local routing protocols. We have tabulated-based routing protocols on either the grounds of scalability, energy consciousness, data aggregation, QoS, and versatility.

References

1. Paul A, Sato T (2017) Localization in wireless sensor networks: a survey on algorithms, measurement techniques, applications and challenges. J Sens Actuator Netw 6(4):24
2. Han G, Xu H, Trung Q, Jiang J, Hara T (2013) Localization algorithms of wireless sensor networks: a survey. Telecommun Syst 52(4):2419–2436
3. Kumar A, Shwe HY, Wong KJ, Chong PHJ (2017) Location-based routing protocols for wireless sensor networks: a survey. Wirel Sens Netw 9:25–72. https://doi.org/10.4236/wsn.2017.91003
4. Hamdi M, Essaddi N, Boudriga N (2008) Energy-efficient routing in wireless sensor networks using probabilistic strategies. In: IEEE communications society in the WCNC 2008 proceedings
5. Jin Z, Ma Y, Su Y (2017) A Q-learning-based delay-aware routing algorithm to extend the lifetime of underwater sensor networks. Sensors 17(7):1660

6. Al Salti F, Alzeidi N, Arafeh BR (2017) EMGGR: an energy efficient multipath grid-based geographic routing protocol for underwater wireless sensor networks. Wireless Netw 23(4):1301–1314
7. Cao N, Wang Y, Ding J, Zhou C, Li Y, Zhang Y, Wang X, Li C, Li H (2017) The comparisons of different location-based routing protocols in wireless sensor networks. In: 2017 IEEE international conference on computational science and engineering (CSE) and IEEE international conference on embedded and ubiquitous computing (EUC)
8. Amiri E, Keshavarz H, Alizadeh M, Zamani M, Khodadadi T (2014) Energy efficient routing in wireless sensor networks based on fuzzy ant colony optimization. Int J Distrib Sens Netw 2014. Hindawi Publishing Corporation
9. Vinutha CB, Nalini N, Veeresh BS (2017) Energy efficient wireless sensor network using neural network based smart sampling and reliable routing protocol. In: IEEE WiSPNET 2017 conference
10. Jafari M, Khotanlou H (2013) A routing algorithm based an ant colony, local search and Fuzzy inference to improve energy consumption in wireless sensor networks. Int J Electr Comput Eng 3:640–650

Design of a Highly Reliable Low Power Stacked Inverter-Based SRAM Cell with Advanced Self-recoverability from Soft Errors

M. Hanumanthu, L. Likhitha, S. Prameela, and G. Pavan Teja Reddy

Abstract Soft error self-recoverability SRAM cell is designed as an extremely reliable SRAM cell. Because the SRAM cell consists of interior nodes which have a unique feedback mechanism and a larger number of access transistors, it is more expensive than the SESRS cell that is a conventional SRAM cell. has several advantages over standard SRAM cell such as: (1) Single event upsets (SNUs) and double event upsets (DNUs) can be self-recovered. (2) It has the potential to reduce electricity consumption usage by 60.2% and silicon area by 23%, when compared to the only SRAM cell on the market that can self-recover from all types of faults. It also has a fourth low-power double-ended stacked inverters which are used in the cell structure of static random access memory for low power consumption. During hold mode and static mode, the power dissipation is even higher which can be decreased by providing inverters with cross-coupling with a low power supply voltage and power gating. In comparison to the 6 T standard SRAM cell, simulation results in the Tanner EDA v16 program using the 65 nm technology library demonstrate a 47.80% reduction in overall power dissipation, a 20.14% reduction in static power dissipation, and an 83% energy retard product has improved. In comparison to the basic 6 T SRAM cell employing the N-curve methodology, the advanced soft error self-recoverability SRAM cell can speed up read approach time by a percent of 61.83 on total, while the proposed SRAM cell has more write ability.

Keywords SNU · DNU · SRAM · Self-recoverability · Soft errors · Static power

1 Introduction

The limiting of power is necessary to make sure the battery's life time, the system's effectiveness, and the data's constant. High integration density and enhanced performance are possible because of aggressive technological scaling used in the manufacture of current advanced SRAM memory. However, because of the voltages

M. Hanumanthu (✉) · L. Likhitha · S. Prameela · G. Pavan Teja Reddy
Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: mhanumanth@gmail.com

405

and capacitances at the node, the amount of charge held on an SRAM cell's node reduces. Soft errors in SRAM memory can cause the data corruption or, in advanced technologies, SRAM circuits to crash.

As a result, hardened designs are essential for obtaining high SRAM reliability in terms of soft errors, as well as developing a low-power, high-efficiency SRAM. The device's feature size has reduced to satisfy the needs of more integration and compact design, resulting in a lower threshold voltage and a thinner gate oxide. It makes the cell structure to waste power even when transistors are turned off, lowering device's complete performance. A thick track that an electron and hole pairs could be created when particle impacts an off state transistor in the SRAM cell. Because of this, the transient pulse which can also be created at the charge collection and recognized at that logic gate's output. If the SET pulse reaches a storage element via downstream combinational logic gates, it may be captured, resulting the storage element retains an invalid value.

The particle, on the other hand, could collide with the transistor which is in off state in the cache element, creating a single node errors or upsets (SNUs). Furthermore, SRAM over all density is growing constantly, while node space is decreasing [1]. One commanding particle might be concurrently influence the two transistors which are off in the cache element can cause multi-node impose collection processes, results in a double node upset (DNU) or double node errors. Indeed, SNU's and DNU's can result in invalid value retention in an SRAM cell. As a result, circuit designers must affect soft mistakes to increase the SRAM cells for protection and criticizing applications and future uses in order to increase the lustiness of SRAM cells. They can preserve themselves from information corruption, implementation or performance problems, and even crashes. Lowering the voltage supply or modifying the threshold voltages can minimize the power utilization of the SRAM cell [2]. The SRAM cell degrades and becomes unstable. The threshold voltage is dropped to improve the device's performance, resulting in a considerable reduction in static power consumption as a result of larger sub-threshold leakage current [3]. Less power approaches like as clasp mode of power grating [4], oscillatory logic [5], and practical ground during the clutch mode [6], for example, it has been proposed to reduce leakage and power dissipation. By lowering the number of transistors, the proposed technology can reduce power consumption.

2   Schematic and Its Normal Operation with Simulation Result

The circuit diagram of the proposed self-recoverability static RAM cell is shown in Fig. 1. The soft errors self-recoverability SRAM cell, as shown in Fig. 1 is made up of 16 transistors that are 8 PMOS transistors from P1 to P8, and 8 NMOS transistors from N1 to N8 are two types. The schematic storage section contains eight PMOS transistors which are P1 to P8 and four NMOS transistors which are N1 to N4. For

access operations, transistor N5 through N8 is employed, with their gate terminals connected to the word line WL. Access transistors N5 to N8 connect internal nodes I1 to I4 to bit line and bit line inverse (BL and BLN), respectively, in the proposed S8P8N cell. The suggested S8P8N cell is depicted in this diagram. When Word Line = 1, the four access transistors will be switched on and allows the write or read access operations. Four access transistors are turned off when WL = 0, and the suggested cell remains the value stored. The proposed S8P8N cell general operations are comparable to the proposed S4P8N cell's general operations. As a result, there are no extensive descriptions here. The results for the proposed S8P8N cell's regular operations are shown below in Fig. 2.

Consider the following scenario to demonstrate the suggested SESRS cell's theory of operation that is the nodes I1 and I3 is equal to 1, and the nodes I2 and I4 are equal to 0. The operations are given as follows. Currently, let's consider the write access operation that is the value of WL = 1, BL = 1, and BLN = 0. In this case, transistors P1, P3, P5, P7, N2, and N4 are OFF, and the remaining transistors are in ON position. Therefore, the output waveforms obtained consist of values of I1 and I3 as 1 and the I2 and I4 as 0. Alternatively, if we change the values, the transistors which are ON in above condition will become OFF, and the OFF transistors will be ON so that the outcomes will be the polar opposite of the first condition.
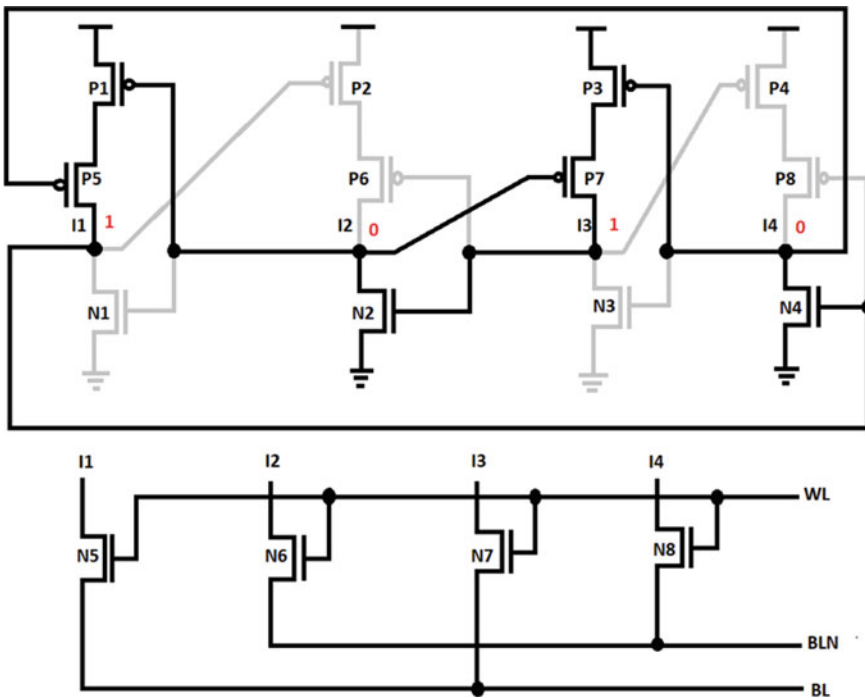
Fig. 1 Schematic of the proposed self-recover soft error SRAM cell
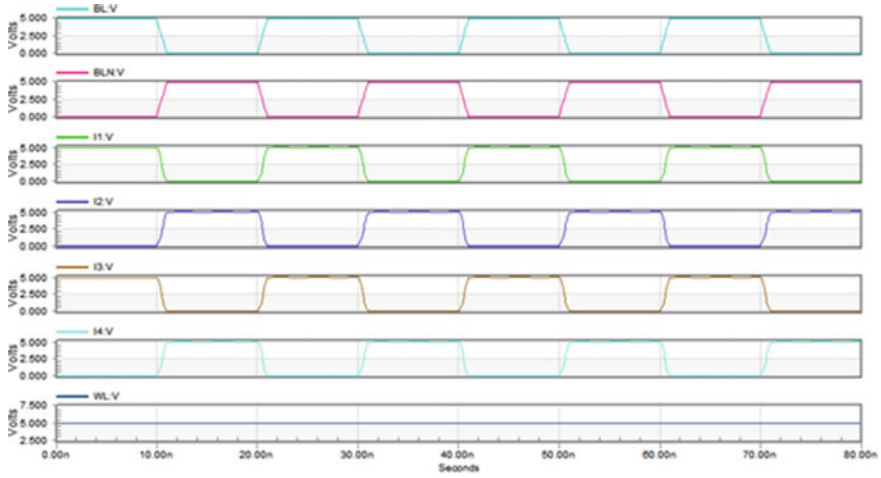
Fig. 2 Simulation results for the proposed self-recover soft error SRAM cell

3 Single Node Upset Self-recovering Principle

It is implemented in the depicted state in Fig. 3. At first, we go for the condition in which I1 node is momentarily altered to 0 from 1 as a result of the SNU. P2, P8 transistors will be ON, and the N4 turns OFF at this time because I1 is flipped.
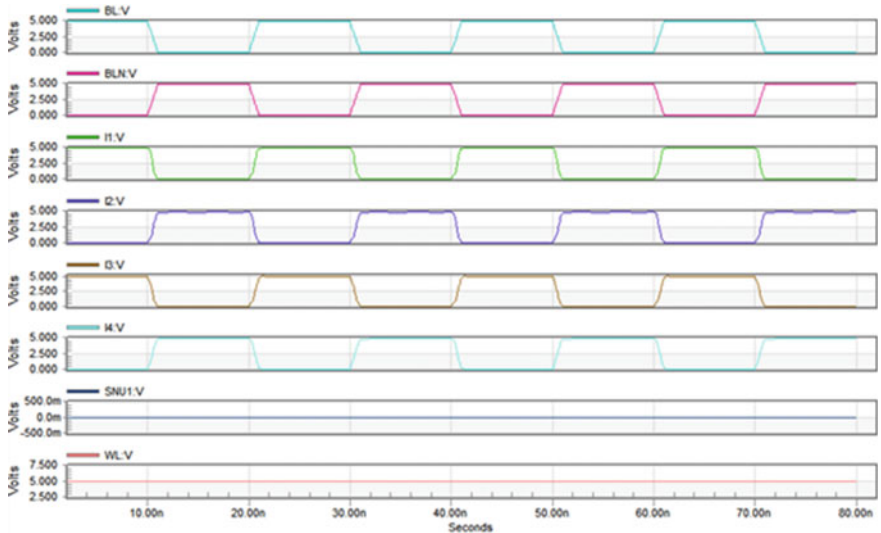
Fig. 3 Simulation results for single event upset self-recovery of the proposed SESRS cell

P4, P6 transistors will be OFF while N2 transistor stays ON since I3 is unaffected, i.e., I3 = 1. As a result, I2 is unaffected (i.e., I2 = 0), and I4 retains its value of 0. P1, P5 transistors remain ON, whereas N1 stays in OFF state since I2 = 0 and I4 = 0. As a result, I1 can come to its normal state, i.e., I1 as 1. As a result, I1 is able to recover from the SNU. The same process is used for the node I3, i.e., when I3 is momentarily adjusted to 0 from its state. Then, we investigate the case when I2 is momentarily changed to 1 from 0 as a result of the SNU. N1 turns on, while P1, P7 will be turned off. So that, I1 is interim converted from 1 to 0; P2 & P8 are momentarily ON, and the N4 is briefly OFF. Transistors P4, P6 continue in OFF state, while N2 is switched ON, because I3 is unaffected by the single node error (it still consists of its value of 1). I2 can thus come to its normal proper value as 0, while I4 retains its earlier value as 0. P1, P5 can come to the normal ON states, and N1 may jumps to OFF state because I2 and I4 are both 0. As a result, I1 can revert to its previous state. To put it in another way, the cell has the ability to recover from the SNU on its own.

4   Double Node Upset Self-recovering Principle

Now take a look at the proposed S8P8N cell's DNU self-recovery principle. There are six critical node pairs in the cell. The scenario of 1 being stored is discussed here. First, examine the case in which is impacted by the DNU, i.e., the I1 momentarily set to 1 and the I2 briefly set to 0. P2, P8, and N1 turn ON when I1 = 0 and I2 = 1, while P1, P7, and N4 become OFF when I1 = 0. Whereas P4 and P6 stay in OFF, while N2 stays in ON, because I3 was unaffected by DNU (I3 still consists of its original accurate value as 1). As a result, I4 retains its prior value as 0; P5 comes into ON, and the I2 reverts to the previous value as 0. So that P1, P7 come to its ON states, while N1 will be into its OFF state, allowing I1 to revert to the right value of 1. That is the S8P8N can self-heal from the DNU. On the other hand, we investigate the condition where the circuit is contrived by DNU that is I3 will currently set to 0, and then, I4 is set to 1. The transistors P4, P6, and N3 will be ON at this time, whereas P3, P5, and N2 will be in OFF state. Because the DNU has no effect on I1, I1 retains its previous accurate value as 1 and P2, P8 transistors remain OFF, and N4 transistor remains ON. As a result, I2 retains its earlier value as 0; P7 stays active, and the node I4 will revert to the previous value as 0 allowing I3 to revert to the value of 1. To put it in another way, the DNU could be self-recovered (Fig. 4).

5   Conclusion

Self-recovery from inconsequential errors A SRAM cell is intended to be extremely trustworthy in its operation. In the hold and static modes, power dissipation is significantly higher than normal; nevertheless, it is possible to lower it by providing inverters
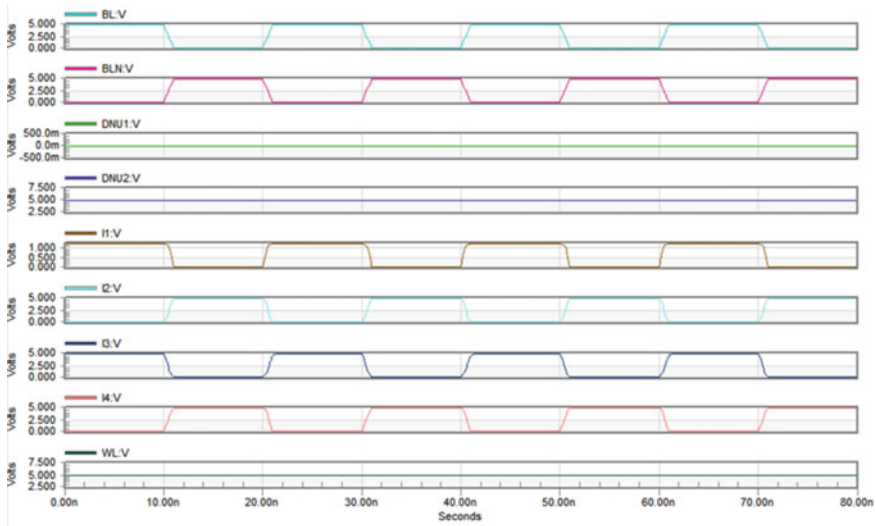
Fig. 4 Simulation results for double event upset self-recovery of the proposed SESRS cell

with cross-coupling and a low power supply voltage, as well as by using power gating. Using the 65 nm technology library, the Tanner EDA v16 software achieved a reduction of 47.80% in overall power usage, which is an encouraging statistic.

References

1. Ebara M, Yamada K, Kojima K et al (2019) Process dependence of soft errors induced by α particles, heavy ions, and high energy neutrons on flip flops in FDSOI. IEEE J Electron Devices Soc 7(1):817–824
2. Nayak D, Acharya DP, Rout PK, Mahapatra KK (2014) Design of low-leakage and high writable proposed SRAM cell structure. In: International conference on electronics and communication systems (ICECS), Coimbatore, 2014, pp 1–5
3. Cai S, Wang W, Yu F et al (2019) Single event transient propagation probabilities analysis for nanometer CMOS circuits. J Electron Test Theory Appl 35(2):163–172
4. Upadhyay P, Mandal A, Kar R, Mandal D, Ghoshal SP (2018) A low power 9T SRAM cell design for both active and passive mode. In: 15th International conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON), Chiang Rai, Thailand, 2018, pp 672–675
5. Narasimham B, Gupta S, Reed D, et al (2018) Scaling trends and bias dependence of the soft error rate of 16 nm and 7 nm FinFET SRAMs. IntReliab Phys Symp 1–4
6. Razavipour G, Afzali-Kusha A, Pedram M (2009) Design and analysis of two low-power sram cell structures. IEEE Trans Very Large Scale Integr (VLSI) Syst 17(10):1551–1555
7. Yan A, Huang Z, Yi M, Xu X, Ouyang Y, Liang H (2017) Doublenode-upset-resilient latch design for nanoscale CMOS technology. IEEE Trans Very Large Scale Integr (VLSI) Syst 25(6):1978–1982
8. Jung I, Kim Y, Lombardi F (2012) A novel sort error hardened 10T SRAM cells for low voltage operation. In: IEEE international Midwest symposium on circuits and systems, pp 714–717

9. Li Y et al (2017) A quatro-based 65-nm flip-flop circuit for soft-error resilience. IEEE Trans Nucl Sci 64(6):1554–1561

10. Zhang J, He Y, Wu X, Zhang B (2018) A Disturb-Free 10T SRAM cell with high read stability and write ability for ultra-low voltage operations. In: IEEE Asia Pacific conference on circuits and systems (APCCAS), Chengdu, 2018, pp 305–308

11. Kobayashi K et al (2014) A low-power and area-efficient radiation-hard redundant flip-flop, DICE ACFF, in a 65 nm thin-BOX FD-SOI. IEEE Trans Nucl Sci 61(4):1881–1888

12. Dang L, Kim J, Chang I (2017) We-Quatro: radiation-hardened SRAM cell with parametric process variation tolerance. IEEE Trans Nucl Sci 64(9):2489–2496

13. Shiyanovskii Y, Rajendran A, Papachristou C (2012) A low power memory cell design for SEU protection against radiation effects. In: Proceedings of NASA/ESA conference on adaptive hardware systems (AHS), Aug 2012, pp 288–295

14. Peng C, Chen Z, Zhang J et al (2017) A radiation harden enhanced quatro (RHEQ) SRAM cell. IEICE Electron Express 14(18):1–12

Image Aesthetic Score Prediction Using Image Captioning

Aakash Pandit, Animesh, Bhuvesh Kumar Gautam, and Ritu Agarwal

Abstract Different kinds of images induce different kinds of stimuli in humans. Certain types of images tend to activate specific parts of our brain. Professional photographers use methods and techniques like rule of thirds, exposure, etc., to click an appealing photograph. Image aesthetic is a partially subjective topic as there are some aspects of the image that are more appealing to the person's eye than the others, and the paper presents a novel technique to generate a typical score of the quality of an image by using the image captioning technique. The model for image captioning model has been trained using convolutional neural network, long short-term memory, recurrent neural networks, and attention layer. After the image caption generation we made, a textual analysis is done using RNN-LSTM, embedding layer, LSTM layer, and sigmoid function, and then the score of the image is predicted for its aesthetic quality.

Keywords Image aesthetic · Convolutional neural network · Long short-term memory · Recurrent neural networks · Attention layer · Embedding layer · Image captioning

1 Introduction

The aesthetic quality of an art piece, like a photograph, concludes in psychological reactions in people. There are various angles that empower a high aesthetic nature of a photograph. As opposed to individuals liking the genuine characteristics of a photograph, they often like different abstract characteristics of a photograph also, for example, regardless of whether the composition is balanced and how the color is dispersed.

Low-level image elements like sharpness, clarity, and saliency that intently identify with human comprehension are named as picture image aesthetic features. It is

A. Pandit (✉) · Animesh · B. K. Gautam · R. Agarwal
Information Technology, Delhi Technological University, Delhi, India
e-mail: pandit.aakash3@gmail.com

Fig. 1 Aesthetic photographs

Fig. 2 Unaesthetic photographs

difficult to plan a metric to evaluate these aesthetic characteristics; however, different investigations show that algorithms can be planned and tuned to anticipate metrics identified with color and composition.

In photography, it typically implies that a picture appeals to the eye. There is something about its composition, subject, and shading that makes us notice it. Similar to beauty, aesthetics are not easy to be characterized in basic words. Everything relies upon the viewer's photographic knowledge, experiences, and preferences. Figure 1 shows some aesthetic photographs which are appealing to the eyes, whereas Fig. 2 shows some unaesthetic photographs.

Photographic artists or specialists purposefully join such properties to frame a bunch of visual guidelines, to catch high-fidelity and appealing pictures to satisfy the audience or want the emotional impact for a huge audience.

There are a few aesthetic principles of photography as follows.

1.1 Composition

- Leading lines—It is the procedure of outlining lines inside a picture to attract the viewer's eye toward that specific point. Since leading lines catch the viewer's concentration, it attracts them to see the magnificence of image.
- Rule of thirds—It is the method involved with dividing a picture into one-thirds, utilizing two vertical and two horizontal lines. This imaginary framework yields nine sections with four points of intersection. Utilizing the rule of thirds makes adjusted creations that are normally satisfying to the eye.
- Symmetry—It is the conventional equilibrium of weight inside a picture. Similar to leading line and *rule of thirds*, it is likewise satisfying to the human eye since symmetry handles the whole image in a balanced way.

1.2 Lighting

- Using available lighting—Trying different things with light does not generally mean utilizing artificial light sources and staging them to make delightful lighting styles. This implies utilizing whatever light is accessible, such that it catches and turns into a textural part of the picture.
- Using shadows creatively— An incredible method for utilizing light to make fascinating pieces is by making shadows. Shadows can make intriguing surfaces and subjects that become a piece of the photograph's composition when it is utilized imaginatively.
- Soft or harsh light—Figuring out which nature of light we like more will likewise assist with developing one's own aesthetic feature. Soft light accomplished falsely with diffusers or normally at Golden Hour can contribute apparently to the pictures, whereas harsh light can assist with making more emotional pictures.

1.3 Color Schemes

- Use colors with intention—Both "color photography" and "black and white photography" have their advantages and disadvantages. It is important to utilize the color intentionally. Purposeful utilization of explicit coloring plans is an incredible method for making an image that is aesthetically pleasing.
- Color theory—Color theory means that the different combinations of colors can have different effects on the viewer, psychologically. So for shooting aesthetical images, the photographer needs to understand the color theory. Stated in section IA, IB and IC are only a few of the many parameters which are there to check the quality of image aesthetic.

2 Related Work

In this section, we discuss some of the state-of-the-art techniques from the literature such as (1) visual attention mechanism, (2) multimodal analysis, and (3) computational approach for image aesthetics score prediction.

2.1 Visual Attention Mechanism

Attention mechanisms started from the examination of the vision of humans. In cognitive science, of the bottlenecks of data processing in the brain, just a small portion

of all apparent data is seen by humans. Roused by this visual attention mechanism, scientists have attempted to track down the model of visual selective attention to reproduce the visual perception process of humans and model the distribution of attention of humans while observing photos and videos. Taking into consideration its widespread applications, a great number models of visual attention have been put forward in the literature, as in [1, 2]. Only recently, the potential success of incorporating visual attention mechanism in image aesthetic prediction has been given more attention by researchers. These methods used recently undermine the process of human perception though they achieve great performance by using saliency to sampling, although these methods achieve impressive performance. While we are observing the images, we give attention to various portions of visual space sequentially in order to gain any important information, and to make an internal representation of the scene, we join the portions of data from multiple fixations. The trial results exhibit that it can accomplish better performance than customary techniques.

2.2 *Multimodal Analysis*

With the fast development of multimedia information, various types of modalities that depict same content can be effortlessly acquired, like sound, pictures, and text. These modalities are interrelated to each other and can give corresponding data to one another. Single-modal methodologies have been outweighed by multimodal methodologies in many works, for example, He et al. in [3] and Bai et al. in [4], for image classification, the authors have joined language and vision to boost up the performance. Multimodal methods have not been used as much in image aesthetic prediction despite having success in many tasks, with a few exceptions like [5, 6]. The only difficult work in multimodal methodology is to combine the multimodal information optimally. Test results exhibit it can accomplish critical improvement on the image aesthetic score prediction tasks.

2.3 *Computational Approach for Image Aesthetic Score Prediction*

The purpose of computational image aesthetics is to design algorithms to perform aesthetic predictions, in a similar way as human beings. In the past two decades, there has been a lot of development in computational image aesthetics prediction, and the credit goes to deep learning algorithms and huge annotated datasets. It has influenced image retrieval, image enhancement, and recommendations in many ways. Many researchers have attempted to tackle the problem of image aesthetic prediction [7–13]. The earlier approaches were based on handcrafted features. In the features here, colorfulness, hue, saturation, and exposure of light [7, 8] are global features

that can be used for all types of photos. Local features such as composition, clarity contrast, dark channel, and geometry should be planned as per the assortment of photograph content [7, 14]. Deep learning networks are often being used nowadays by researchers for image aesthetic quality assessment. Few fundamental works in present-day computational aesthetics prediction were put forward by Datta et al. [8] and Ke et al. [15] in 2006. Kucer et al. [12] with the deep learned features consolidate the hand-designed features to lift up the performance. Kao et al. [11], for the extra supervision, used tags on images to predict the image aesthetics and put forward a deep multitask network. The methods pose attention on encoding the composition (global) and finer details (local). The local view is addressed by cropped patch, and global view is addressed with a distorted image.

3 Proposed Methodology

To check the aesthetic quality of an image, we have used deep learning models and have obtained a decent quality checker.

3.1 CNN

Image processing and identification, which mainly deals with the pixel information, is done by convolutional neural networks (CNN), which are a type of artificial neural network (ANN). A CNN uses a framework that decreases the processing necessities like a multilayer perceptron. The layers of a CNN consist of three layers which are input, output, and a hidden layer. They help in incorporating different pooling layers, convolutional layers, normalization, and fully connected layers.

3.2 Caption Generator

To increase the quality of our prediction, we check the quality of images using the content inside the image, and for obtaining that, we have extracted an image caption for every image, and then from that generated caption, we have predicted the aesthetic quality of the image. For image caption generator, we have used [16] which follows the model structure as stated below.

- Encoder: CNN—The model takes a raw picture and creates an encoded subtitle. We have utilized a CNN for extracting the feature vectors. We have taken features from a lower convolutional layer which permits the decoder to specifically focus on specific pieces of a picture by choosing a part of all the feature vectors.

- Decoder: LSTM—Long short-term memory (LSTM) networks are a kind of recurrent neural networks (RNN) used in prediction problems of sequencing for learning order dependence. We utilize an LSTM network that creates a caption by producing a single word at each time step based on a context vector, the recently generated words, and the previously hidden states. Figure 3 shows an LSTM cell.
- RNN—RNNs are an incredible and vigorous sort of neural networks and consist of a very promising algorithm because they have internal memory. Due to their internal memory, RNNs can recollect significant things about the information they obtain, which permits them to have a precise prediction of what is coming in the future.
- Deterministic soft attention layer—Attention is a procedure that imitates intellectual attention. The impact upgrades the significant pieces of the input information and fades the rest. Learning stochastic attention requires testing the attention area s_t each time, rather we can make the assumption for the context vector z_t and plan a model which is deterministically attentive.

$$E_{p(s_t\|a)}[\hat{z}_t] = \sum_{i=1}^{L} \alpha_{t,i} \mathbf{a}_i \tag{1}$$

The entire model is smooth and differentiable under deterministic attention, so using backpropagation to learn end-to-end is trivial. So, the model is prepared end-to-end by limiting the negative log probability:

$$L_d = -\log(P(\mathbf{y}\|\mathbf{x})) + \lambda \sum_{i}^{L}(1 - \sum_{i}^{C} \alpha_{t_i}) \tag{2}$$

3.3 Caption Analysis

We analyze the caption of image using RNN-LSTM model for generating the score of the image. We use three layers of network as shown in Fig. 4:

- Embedding layer—We have too many words in our vocabulary, so it is computationally very expensive to do one-hot encoding of these many classes, so we add an embedding layer as the initial layer. We utilize this layer as a lookup table instead of one-hot encoding.
- LSTM layer—Long short-term memory layer takes care of results obtained from the previous layers as it stores two types of memory, long-term memory and short-term memory, within itself. It has four gates: learn gate, forgot gate, remember gate, and use gate; an LSTM cell is shown in Fig. 3. The input data goes in this layer, and the layer with the help of previous and current information predicts the result.

Fig. 3 LSTM cell [16]
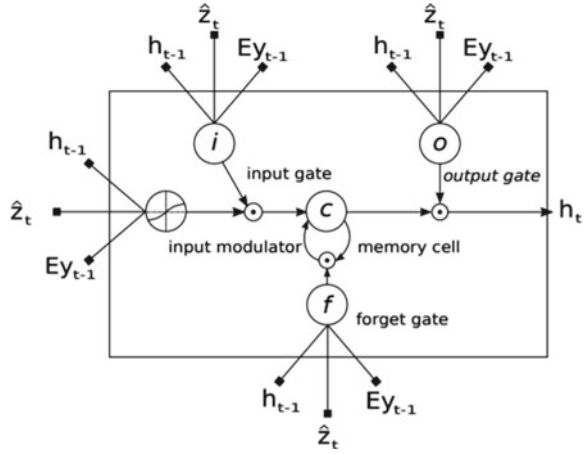
Fig. 4 Layers of our model
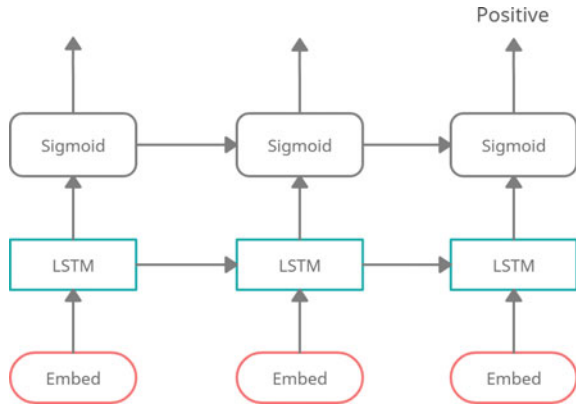
- Sigmoid function—A single sigmoid output gives a probability between 0 and 1, which we use in the multi-cross-entropy loss as the loss function.

$$S(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

4 Experimental Analysis

4.1 Training Dataset

We have used aesthetics and attribute database (AADB) given by Kong et al. [17]. It consists of 8000 images with eight attributes (vivid color, rule of thirds, object

Table 1 Comparison of AADB dataset with AVA and Photo.Net dataset

	AADB	AVA	Photo.Net
Rater's ID	Yes	No	No
Real photos (All of them)	Yes	No	Yes
Attribute label	Yes	Yes	No
Score distribution	Yes	Yes	Yes

emphasis, color harmony, good lighting, shallow depth of field, interesting content, balancing element) having overall scores of aesthetic quality and also having binary labels of them which have been rated by five Amazon Mechanical Turk (AMT) workers. The AADB and AVA give a bigger scope, more extensive score circulation, more extravagant semantic, and style characteristic comments than Photo.Net. AADB contains a significantly more adjusted distribution of visual symbolism of genuine scenes downloaded from Flickr. However, the number of images is less compared to other datasets, and the attribute tag is binary in the AADB dataset (high or low aesthetic). Table 1 shows the comparison of AADB dataset with AVA and Photo.Net dataset.

This dataset contains rater identities and informative attributes along with the score distributions of images. These explanations empower us to concentrate on the utilization of people's rating on the quality of image aesthetic for training our model and investigate how the trained model performs contrasted with individual human raters.

4.2 Score Prediction with Caption Generation

- Training dataset for caption generation—Here we have used the MS-COCO dataset to train our caption generation model. The MS-COCO dataset is a large-scale captioning, object detection, and segmentation dataset published by Microsoft [18]. It is a popular dataset among the computer vision pioneers in the machine learning community. As the main task of computer vision is to understand the scene and conclude what is going on in a scene and then coming up with a semantic description. The state-of-the-art MS-COCO data is suitable for this task of image captioning. There are more than 82,000 images in this dataset, with each image having no less than five captions with different annotations. The dataset has about 91 categories of objects. COCO has less categories and more instances per category. We are using 20,000 images with 30,000 captions, where each image is having multiple captions from the MS-COCO dataset. We create 80–20 split randomly of training and validation sets. So there are 16,000 images for training and 4000 for testing. As MS-COCO dataset is very large and reliable, we used it to train our model.
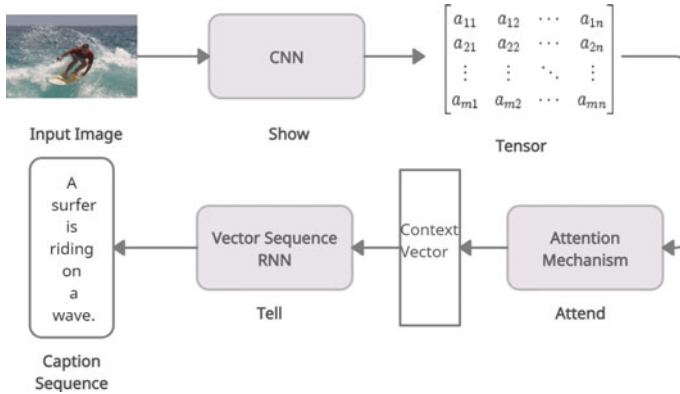
Fig. 5 Caption generation model

- Caption generation—Images are resized to 299 px * 299 px and passed to Inception V3 model which classifies the images, and features are extracted using the last convolutional layer of the Inception V3 model. For processing the captions, we first tokenize them, which gives us the vocabulary of all the words which are unique in the dataset, and then we limit the size of the vocabulary up to 5000 words only to save the memory and then replace other words with "UNK"(unknown) token. After that we create mappings of index-to-word and word-to-index and then pad zeroes at the end of each sequence to make all the sequences of same length, where the length would be of the longest sequence available. Attention layer allows our model to focus on the parts of the image which are of prominence in generating the caption and gives better results for our caption generator model. After extracting the vector from lower convolutional layer of the Inception V3 model, we pass this vector through CNN encoder which has a single fully connected layer. The outputs of the encoder, hidden state (initialized to zero), and start token are passed to the decoder. Decoder returns its hidden state and the predictions of caption. Then the hidden state is returned back, and the predictions are utilized for loss calculations. The RNN is then used to analyze the image to predict the next word. Figure 5 shows the caption generation model, and Fig. 6 shows an example of a caption generated by our model.
- Generating aesthetic score from captions—The file of captions generated by our caption generator model is used as an input for our next model which generates an aesthetic score from the captions. Steps for doing the same are as follows:

1. The first step of this model encodes each word with a number using embedding layer as the first layer. This step converts our data suitable for feeding into a neural network. We remove all punctuations and emojis and split the text into the words in a single line.
2. We utilize the spacy and regular expression to take out only the textual data for the process of lemmatization and tokenization of each word of our caption.
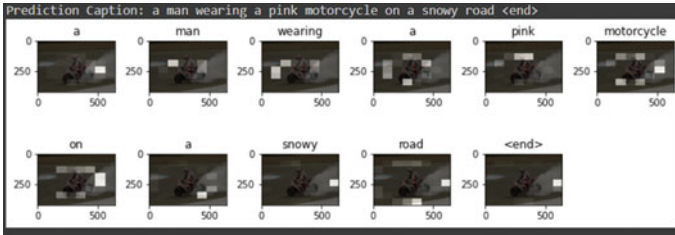
Fig. 6 Image caption generated through attention mechanism

3. We count all the distinct words in our vocabulary and sort the number according to their occurrence and map them starting from one.

Our labels are also integers, from one to five. So, we have used one-hot encoding here. We convert the captions to the same length for accurate results and less computation. We fix a specific input length for our captions, and then we convert all other captions to that specific length only. To do this task, we followed two steps:

(a) We delete extremely small and extremely large captions.
(b) For the small captions, we pad zeroes at the end, and for large captions, we truncate it from the end.

For the above step, we choose the fixed length to be twenty in this project. So, whenever a caption size is less than twenty, we pad zero at the end, and if its size is greater than twenty, we truncate it from the end. Now our data is of fixed length, so now we divide it into the testing, training, and validation data. While training our model, for changing the attributes/weights of RNN we use Adam optimizer, as it is robust, and also it takes momentum into consideration of the updated gradients. We use the LSTM layer for predicting the result with the help of previous information and the current information and use a sigmoid function for entropy loss calculation. Figure 7 shows the flow of our textual analysis model which generates a score from the captions.

4.3 General Model

So, in our general model, the image is taken as an input for the caption generator model. This model generates a caption from the image, which gives a description of the image. Then the generated caption of the image is provided as an input for the text analysis model, which then predicts the aesthetic quality of the image by generating a score on the scale of 1–5. Figure 8 shows the flowchart of our general model.
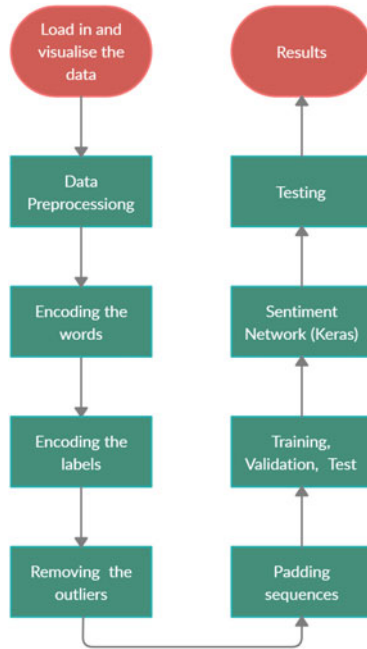
Fig. 7 Flowchart for textual analysis model
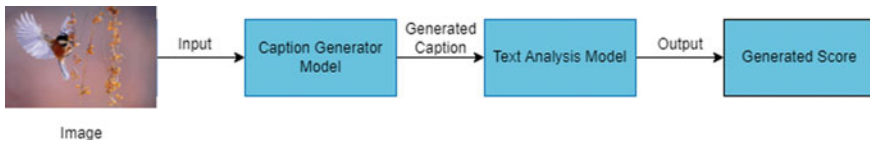
Fig. 8 General model

5 Results

First we used the CNN model inside the images to test the accuracy of prediction of the quality of the images. After testing, we got a low accuracy of 28.72%. Because of this low accuracy, we tried to make different type CNN models and combined the results of all the CNN models to generate the score of the image. With this, the accuracy improved to around 30–35% accuracy.

In our text analysis model, we made a vocabulary of words, and it consisted of a total of 1687 words and after which we got after this process was of around ∼ 40%. We tested the model using dataset of 800 images, and the model gave correct accuracy as given in Table 2.

Table 2 Summary of the scores generated by our model

Score	1	2	3	4	5
Predicted count	3	10	276	24	6
Groundtruth count	40	127	357	188	88

Table 2 shows the number of predicted images which have same scores as of ground truth images and are distributed using scores ranging from 1 to 5. The extreme low and high scores, like 1 and 5, are predicted less accurately because of less sample size of these scores.

6 Conclusion

In this paper, we calculate image aesthetic score by first generating the captions for the image and then use it to predict the aesthetic score. We predict the aesthetic score of an image on a scale of 1–5, and while taking into consideration the subjectivity of the task, our model shows promising results. Future work may be about using photo critique captions to enhance the performance. Image aesthetics is a partially subjective topic as there are some aspects of the image that are more appealing to the person's eye than the others, and the paper presents a novel technique to generate a typical score about the quality of an image by using the image captioning technique. The image captioning model has been trained using convolutional neural network, long short-term memory, recurrent neural networks, and attention layer, and we achieved the accuracy of $\sim 40\%$.

References

1. Obeso AM, Benois-Pineau J, Vazquez MSG, Acosta AAR (2019) Forward-backward visual saliency propagation in deep NNs vs internal attentional mechanisms. In: 2019 9th International conference on image processing theory, tools and applications, IPTA 2019. https://doi.org/10.1109/IPTA.2019.8936125
2. Mnih V, Heess N, Graves A, Kavukcuoglu K (2014) Recurrent models of visual attention. Adv Neural Inf Process Syst 3(January):2204–2212
3. He X, Peng Y (2017) Fine-grained image classification via combining vision and language. In: Proceedings—30th IEEE conference on computer vision and pattern recognition, CVPR 2017, vol 2017-Janua, pp 7332–7340. https://doi.org/10.1109/CVPR.2017.775
4. Bai X, Yang M, Lyu P, Xu Y, Luo J (2018) Integrating scene text and visual appearance for fine-grained image classification. IEEE Access 6:66322–66335. https://doi.org/10.1109/ACCESS.2018.2878899
5. Yu Z, Yu J, Xiang C, Fan J, Tao D (2018) Beyond bilinear: generalized multimodal factorized high-order pooling for visual question answering. IEEE Trans Neural Netw Learn Syst 29(12):5947–5959. https://doi.org/10.1109/TNNLS.2018.2817340

6. Zhou Y, Lu X, Zhang J, Wang JZ (2016) Joint image and text representation for aesthetics analysis. In: MM 2016—Proceedings of the 2016 ACM multimedia conference, pp 262–266. https://doi.org/10.1145/2964284.2967223

7. Tang X, Luo W, Wang X (2013) Content-based photo quality assessment. IEEE Trans Multimedia 15(8):1930–1943. https://doi.org/10.1109/TMM.2013.2269899

8. Datta R, Joshi D, Li J, Wang JZ (2006) Studying aesthetics in photographic images using a computational approach. In: Lecture Notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 3953. LNCS, pp 288–301. https://doi.org/10.1007/11744078_23

9. Guo L, Xiong Y, Huang Q, Li X (2014) Image esthetic assessment using both hand-crafting and semantic features. Neurocomputing 143:14–26. https://doi.org/10.1016/j.neucom.2014.06.029

10. Nishiyama M, Okabe T, Sato I, Sato Y (2011) Aesthetic quality classification of photographs based on color harmony. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp 33–40. https://doi.org/10.1109/CVPR.2011.5995539

11. Kao DY, He R, Huang K (2017) Deep aesthetic quality assessment with semantic information. IEEE Trans Image Process 26(3):1482–1495. https://doi.org/10.1109/TIP.2017.2651399

12. Kucer M, Member S, Loui AC, Messinger DW (2018) For predicting image aesthetics. 27(10):5100–5112

13. Chen Y, Hu Y, Zhang L, Li P, Zhang C (2018) Engineering deep representations for modeling aesthetic perception. IEEE Trans Cybern 48(11):3092–3104. https://doi.org/10.1109/TCYB.2017.2758350

14. Luo Y, Tang X (2008) Photo and video quality evaluation. Quality 8(08):386–399. [Online]. Available: http://portal.acm.org/popBibTex.cfm?id=1478204&ids=SERIES9612.147-8172.1478182.1478204&types=series.proceeding.section.article&reqtype=a-rticle&coll=GUIDE&dl=GUIDE&CFID=91452395&CFTOKEN=97546522

15. Ke Y, Tang X, Jing F (2006) The design of high-level features for photo quality assessment. Proceedings of the IEEE computer society conference on computer vision and pattern recognition 1:419–426. https://doi.org/10.1109/CVPR.2006.303

16. Xu K et al (2015) Show, attend and tell: neural image caption generation with visual attention. In: 32nd International conference on machine learning, ICML 2015, vol 3, pp 2048–2057

17. Kong S, Shen X, Lin Z, Mech R, Fowlkes C (2016) Photo aesthetics ranking network with attributes and content adaptation. In: Lecture Notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 9905. LNCS, pp 662–679. https://doi.org/10.1007/978-3-319-46448-0_40

18. Lin TY et al (2014) Microsoft COCO: common objects in context. In: Lecture Notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 8693. LNCS, no. PART 5, pp 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Pedestrian Crossing Signal Detection System for the Visually Impaired

Swati Shilaskar, Shubhankar Kalekar, Advait Kamathe, Neeraja Khire, Shripad Bhatlawande, and Jyoti Madake

Abstract Navigating in outdoor environments can be a challenge for the visually impaired, especially given the increase of vehicular activity on the streets. It is not wise to rely solely on the people involved in the scenario to help the visually impaired person cross the street safely given the number of accidents that happen due to irresponsible driving. Ideas to tackle these problems have been implemented before, using both machine learning and deep learning techniques. Several papers also employ a variety of sensors like proximity sensors, ultrasonic sensors, etc., in order to get relevant feedback in analog format from the surroundings. Camera is one such sensor that can be used to sense the surroundings in order to get relevant digital input. This paper proposes a computer vision-based technique to use such digital input and process it accordingly in order to help the visually challenged tackle the problem. Simple machine learning solutions like SIFT (for feature extraction) are used. Comparison of different classifiers like SVM, K-means, and decision trees has been done to identify the best classifier for a given requirement. Use of simple and efficient methods can be conducive for deployment in real-time systems. Proposed system has a maximum accuracy of 86%.

S. Shilaskar (✉) · S. Kalekar · A. Kamathe · N. Khire · S. Bhatlawande · J. Madake
Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of
Technology, Pune 411037, India
e-mail: swati.shilaskar@vit.edu

S. Kalekar
e-mail: shubhankar.kalekar19@vit.edu

A. Kamathe
e-mail: advait.kamathe19@vit.edu

N. Khire
e-mail: neeraja.khire19@vit.edu

S. Bhatlawande
e-mail: shripad.bhatlawande@vit.edu

J. Madake
e-mail: jyoti.madake@vit.edu

1 Introduction

Being visually impaired can be the toughest test for anyone in the world. Not being able to see things, recognize faces and elements around is a challenge for them. Many people in the world suffer from this disorder. As of 2021, 33 million people live with avoidable blindness and more than 250 million [1] have moderate-to-severe impairment. The rise in number is solely due to the lack or unavailability of good medical care.

Visual impairment not only restricts the accessibility to new things, but also takes away the independence of these people. One such major challenge that they might face daily is to cross the roads without external help. Various technologies for detecting crosswalks and vehicles have been developed. Most common solution for this aid includes a sound signal associated with the specific state of the signal. Due to its bulky and costly hardware and maintenance, implementation is not feasible. The American council of the Blind stated that the "cuckoo-chirp" type sound installed at the intersections might lead to more accidents due to unclear identification which might create chaos [2]. Using multiple tones for multiple directions was the solution for this problem, but that makes it difficult for the pedestrians to remember and correlate the tones properly which makes this system risky.

Various techniques for detection of traffic lights have been implemented world wide. But, very little research has been done that specifically targets pedestrian traffic lights. Various state-of-the-art (SOTA) deep learning models have been developed to find and recognize the pedestrian signal, but there is seen a gap in detecting them using standalone computer vision techniques. Hence, the main goal of this work was to present a solution to detect the pedestrian traffic signal and recognize its current state so that the people destitute of vision can cross the roads safely and independently. The current SOTA systems in localization, detection and categorization have been surveyed and studied in this paper.

The main contributions of this work are stated as below: (1) Creating a labeled dataset for the purpose of training of the model. The dataset included 3000 images for each class—Positive and Negative. (2) Comparison of various classifiers and techniques like, support vector machines (SVM), K-Means clustering, decision tree, etc. (3) A thorough analysis of the results to prove that the approach was feasible and worthy of implementation and thus be used by the people who need such aid. The flow of the paper is as follows. A literature survey of the existing techniques and methods has been done in upcoming section. The methods and various techniques used to detect and classify the signals have been discussed in methodology section. The analysis regarding the feasibility of implementing this approach for the classification of pedestrian signals find its mention in the results section. Finally, the conclusion along with further work that can be done are included.

2 Literature Review

A large amount of literature is available for this domain of computer vision-based aids for the blind. Machine learning and image processing based approaches are lightweight and useful in real time while deep learning based approaches are accurate, but require high computational resources.

The approach by Kaluwahandi et al. [3] proposes a portable navigation system and an image processing-based approach. They have also detected crosswalks by dividing the image into two parts and accumulating the brightness of both parts in order to determine which part contains the white pedestrian crossing. They also detect traffic control signs using their shapes based on a binarization approach. Again, pedestrian or traffic light detection is not included in this approach.

Chen and Huang [4] have used color extraction methods for selecting candidate regions of interest which are further passed to a classifier for detection of the presence of traffic lights. The classifier is trained on images of traffic lights of different colors as well as having arrows pointing in various directions. Since a traffic light is simply a blob of color and has no other unique features, it is difficult to identify its presence. To solve this problem, they have assumed that traffic lights will always appear in a holder containing 3 or 4 lamps that are vertically aligned.

Suda et al. [5] have used Haar-like features to detect the traffic signal. They propose a detection method using high brightness pedestrian light against its background as the main feature for detection. Apart from that, they have made use of template-matching algorithms for detecting crosswalks.

Similar difficulties are recognized by Wonghabut et al. [6] and have made use of the (Hue-Saturation-Value (HSV) model. Given the weather conditions, the lights may be dimmed, have low brightness and so on. These will mainly affect the Saturation and Value elements and not the Hue. Hence, if the Hue matches, the color can then be identified as belonging to the traffic signal.

Lee et al. [7] have used Haar-like features for training SVM for detection. Yang et al. [8] have combined semantic segmentation with deep learning to help visually impaired people navigate intersections in general. This includes the detections of all the elements of a typical intersection scene including zebra-crossings, pedestrian lights, vehicles, and so on. They have used the fact that crossings and pedestrian lights often appear side-by-side as a context-cue.

Ghilardi et al. [2] have trained their model on the Pedestrian Traffic Detection (PTLD) dataset to identify go and stop conditions as well as when the signal is switched off. They have identified their region of interest by extracting features through neural networks.

The method used by Ying et al. [9] detects circular and rectangular shapes along with analysis of the variation in intensity of the different pixels for detecting traffic lights. It also uses brightness normalization to account for the continuously changing lighting conditions outside. The rectangular frame of the traffic lights is divided and based on which part is lit up, the color is seen as either red or green.

An interesting technique for detecting traffic lights is shown by Mascetti et al. [10] is based on the fact that traffic lights usually have a high luminosity and are surrounded by areas with low light. Hence, they have used a range filter defined using the HSV model to identify pixels (of a particular color) with high luminosity.

Omachi and Omachi [11] have used color space conversion, followed by ROI and edge detection. da Silva Soares et al. [12] use removal of non-circular images followed by morphological filtering and ROI recognition for detecting and traffic lights. Feature extraction is done using texture classifiers and classification is done by neural networks.

Ozcelik et al. [13] had encountered a problem in the detection of traffic lights due to the separate led lights present in the signal. They experimented with normalized box filters, Gaussian filters and median filters to conclude that median filters work best for these types of images. This is because they give good results without blurring the edges too much. Traffic lights may appear as numbers, hand gestures (stop–go), human figures, etc. Hence, Pongseesai and Chamnongthai [14] have applied feature extraction tools to their images.

In order to account for the variation used fuzzy logic to determine whether to take it as a cue to move on or not. Al-Nabulsi et al. [15] have created a traffic light detection system for the color blind. Their method uses 22 different templates of traffic signals with different shapes and cases. These are matched with the ROI which is captured by cropping up the acquired images. They have applied median and linear Kalman filters to enhance the colors of the lights and make the identification process easier. While this is a good approach, template matching can often detect similar unrelated shapes in an image unnecessarily complicating the execution process.

Wu et al. [16] have a vision-based approach to the problem of pedestrian lights detection. They faced the problem of detection of multiple pedestrian lights, vehicle tail lights and misclassification due to different shades of the signals as seen in different kinds of lighting conditions. To combat these problems in the best possible way, they have used a color filter followed by a size and background filter which try to eliminate any existing false positives. The aspect-ratio of the candidate regions is also determined and a score is calculated to determine whether the candidate is really a pedestrian light signal or not.

Sooksatra and Kondo [17] have used fast radial symmetry transform for detection of traffic lights. This has been developed specially to detect circular objects with significantly high or low brightness levels based on radius. The CIELab Color Model is used. This is less dependent on lighting conditions.

Muslu and Bolat [18] have performed vehicle tail light detection. They have also made use of red and white HSV thresholds to find places where red is adjacent to white in order to detect when brakes are applied by the vehicle ahead and brake lights are lit up. They have also performed lamp edge detection by calculating local maxima using the canny edge detector.

Binangkit and Widyantoro [19] have made a summary of all the different techniques that can be used for color and edge detection, as well as circular Hough transform and so on. They have also studied different classifiers like SVM, ANN, and random forest. A sensor-based approach is employed by Mahendran et al. [20] which uses several different techniques to identify various parts of the scene including obstacles as well as other objects. A smart stereo vision sensor is employed and 3D-point cloud processing is used to map the surroundings. The system includes Intel's NCS2 which has the ability to perform the high-end tasks required. One of the obvious limitations as stated by them is that the Faster R-CNN that they have employed has high prediction time which makes the process timing consuming making it unsuitable for real-time situations.

Saleh et al. [21] uses DeepLabV3+ for semantic segmentation of the image using deep learning. DeepLabv3 uses dilated convolutions (a mask with holes) so that it can keep track of the bigger picture while still requiring low computation. This was chosen for its low processing time as well as less amount of memory used and a predetermined set of classes that can be selected for training.

Mallikarjuna et al. [22] present an example of the deployment of a visual aid system on Raspberry Pi using an IoT based approach for collecting sensor data from the Raspberry Pi Camera for real-time processing. However, their approach is deep learning based.

Hsieh et al. [23] have proposed a wearable guiding system taking feedback from an RGB-D camera for enabling visually-impaired people to easily navigate in outdoor environments. For training the system they have collected data of sidewalks, crosswalks, stairs and roads and trained a FAST-SCNN model. The goal is to find a walkable area for the user. This is done by dividing the image into seven parts and analyzing them for features of crosswalks or sidewalks. It is safe for the user to walk where they are detected. Based on which segments the crosswalk/sidewalk appears in, the safest direction to navigate is determined.

Tapu et al. [24] have conducted a state-of-the-art survey of monocular camera, Stereo camera and RGB-D camera-based wearable, portable assistive systems developed for the visually impaired. They conclude that while computer vision and machine learning based approaches have improved over the past years, they are yet to reach human-like abilities for understanding the environment.

To mention a few street-crossing aids, Ivanchenko et al. [25] try to extract straight lines from a test image. The algorithm then tries to combine them using a segmentation-based approach to create a crossing-like pattern. This is done using a simple smart-phone camera.

Tian [26] has used Hough transform to detect a group of parallel lines in front of a blind user. These can be either a staircase or a crosswalk. They have then used an RGB-D sensor for depth estimation to differentiate between staircases and crosswalks. References [27–34] proposed different sensor-based systems to aid the visually impaired. After studying all the different approaches, a computer vision and machine learning based approach seemed better for easy deployment as well as from the point of view of resource utilization.

3 Methodology

The proposed system acquires the images of the surrounding using a camera and detects the presence of the pedestrian light signal. On detection and recognition of the color of the signal, it will provide audio feedback accordingly shown in Fig. 1.

3.1 Dataset Compilation

Majority of the images in the dataset were compiled from existing datasets like PTLD [2], Pedestrian Augmented Traffic Light Dataset [35], and LISA Traffic Light Dataset [36] on the Internet. Along with this some images were captured by Samsung Galaxy M40, and Samsung Galaxy M30s in Pune, Maharashtra, India. Table 1 shows the distribution of the images into classes red, green (positive), and negative.

As a part of preprocessing (refer Fig. 2), each image is resized to 140×215 for uniformity as well as to reduce the size of the data. Brightness normalization is performed on the images. The primary reason for this is that the images consist of lights and can have high contrast levels against their background.

The images are then filtered using the median filter. The reason for applying the filter is that pedestrian lights do not appear as a continuous color due to the spaces

Fig. 1 System diagram
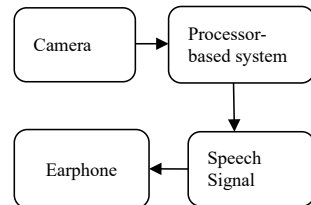
Table 1 Distribution of images in the dataset

Red	Green	Negative
1500	1500	3000

Fig. 2 Image preprocessing
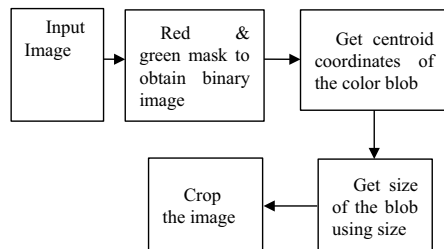
Table 2 HSV color parameters

	Hue	Saturation	Value
Red	$1 \leq H \leq 15$	$35 \leq S \leq 255$	$140 \leq V \leq 255$
Green	$41 \leq H \leq 68$	$128 \leq S \leq 255$	$128 \leq V \leq 255$

between the LEDs. Blurring blends the individual lights into a continuous color patch. This will make it easier for the classifier to detect its presence.

Color segmentation is performed on these images to extract patches of bright red/green color by passing the HSV color parameters (refer Table 2) of those shades. These will be the red and green signals, i.e., the regions of interest (ROI). These ROI are then extracted from the original images by cropping the extra area. Some images obtained after cropping contained patches of red and green that were not signals. Such images were removed from the data manually. Rest of the images are converted to grayscale to further reduce the computing costs.

The final data used for training will therefore consist of close-up, grayscale images of pedestrian signals. Thus, the dataset will have a positive class with images in which the pedestrian signal is present and a negative class in which the signal will be absent.

The dataset was split into five parts: four parts were used for training and one for testing. Five different machine learning based models are trained, tested, and compared. Five-fold cross validation is performed on the classifier with the highest accuracy. Finally, 100 new samples of each class are used for further validation.

3.2 Feature Extraction and Feature Vector Compilation

Scale Invariant Feature Transform (SIFT) finds the scale of the image in which the feature will get the greatest response. It is also unaffected by changes in perspective. This is useful for detecting pedestrian signals from different distances and angles.

SIFT is also invariant to changes in illumination. The process of feature extraction involves the calculation of gradients. The gradient is computed using the difference in various pixel intensities. Hence, if the brightness changes equally everywhere, this does not affect the process of feature extraction. This makes it useful for detecting bright as well as dim lights. The difference occurs usually due to the changes in the surrounding lighting and time of the day in which the image is captured. These changes usually affect all pixels at a time and are uniform throughout the image. SIFT first divides its regions of interest (key-points) in the image into parts of 16 × 16 and further divides these into parts of 4 × 4. Considering that eight vector directions are used to produce the gradient vector, a feature vector of 2,266,128 rows × 128 columns size was obtained.

Fig. 3 Data preprocessing

Input Image (140 x 215)	Color Segmentation to obtained cropped image	SIFT Feature vector obtained: 2266128 rows × 128 columns
PCA output data: 24986 rows x 5 columns	K-means output data: 24986 rows x 6 columns	
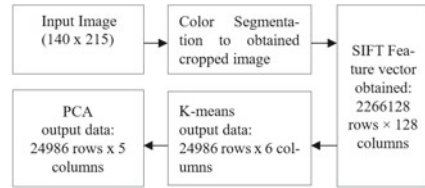
3.3 Dimensionality Reduction

K-means clustering is an unsupervised algorithm which separates the given data into several clusters. The elbow method gave 5 as the optimum number of clusters. Applying K-means reduced the data size to 24,986 rows × 6 columns. Principal component analysis (PCA) condenses high-dimensional data to reduce it to a few major components which represent a large part of the variance in the data. Applying PCA reduced the data size to 24,986 rows × 4 columns. Figure 3 represents the entire data preprocessing pipeline of the proposed system.

3.4 Classification and Detection of Pedestrian Crossing Light

Logistic regression is primarily used for binary classification where the dependent variable is categorical (e.g., positive or negative). It will take the data from the last step and calculate the probability that the image contains a pedestrian signal or not. Support vector machines (SVM) gives the optimum hyperplane to segregate multidimensional data into classes so that any new data given to the model can be categorized easily. Support vectors are the points positioned closest to this decision boundary that affect it the most. SVM tries to keep the margin between the boundary and the vectors as large as possible.

SVM linear kernel assumes that a simple linear boundary can separate the data into the required classes. Polynomial kernel gives an output hyperplane that is nonlinear. This is usually avoided in cases where there can be chances of overfitting. Radial basis function (RBF) kernel draws normal curves around the data points. The decision boundary is defined where the sum of these curves reaches a threshold value. This kernel is not affected by translations in the feature. This is achieved because it works on the difference in the two vectors. This will be useful in case our features have a tendency to vary linearly in magnitude. The linear hyperplane equation is given in Eq. 1.

$$f(x) = B(0) + \mathrm{sum}(a_i * (x, x_i)) \tag{1}$$

where x is input and x_i is support vector. $B(0)$ and a_i are estimated from training data.

K-Nearest Neighbors (KNN) is a supervised algorithm for classifying new data into pre-formed clusters. It selects K nearest points of the query data point and then classifies it according to the most frequently occurring label. A decision tree is a tree which follows a step-by-step process, branching at every node for a given outcome. Each node is a test which eventually leads us to the final outcome. Random forest is a supervised machine learning algorithm which uses multiple decision trees to predict the output. The predictions Y of n trees, each with its own weight function M_i, averaged in a random forest is given in Eq. 2.

$$Y = 1/n \sum_{i=1}^{n} \sum_{j=1}^{m} M_i(x_j, x') y_j \qquad (2)$$

After training, the classifier is used to perform prediction on the test set. Algorithm 1 explains the logic behind classification and color recognition.

Algorithm 1 System Implementation

Input An input image
Output Signal Indication
1. red and green masks are applied to the image
2. **if** red is detected
3.   color = "red"
4. **if** green is detected
5.   color = "green"
6. image is cropped to extract ROI
7. SIFT o ROI to get feature descriptors
8. prediction = classifier.predict (descriptors)
9. **if** prediction is positive:
10.   output = colour + "signal"
11. **else**
12.   output = "no signal"

4 Results

Figure 4 shows the outcome of the HSV color detection and localization and identification of green signal. Similarly, Fig. 5 is the result for the red signal. It is evident that the second figure is difficult to identify even for a visually abled person. This approach can help even a visually impaired person identify it correctly.

Performance Evaluation. Machine learning based models are trained on the dataset mentioned earlier. Performance evaluation of these models from classification chart (see Figs. 6 and 7) and classification report has been mentioned (refer Table 3).

Fig. 4 Green signal detected on a signal input
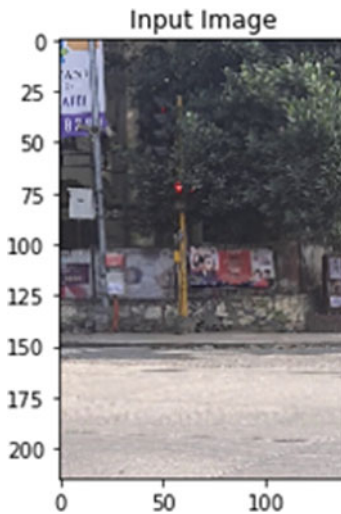
Fig. 5 Red signal detected on a signal input

5 Conclusion

In this paper, a solution to overcome the challenge faced by sight impaired people to cross the roads at the junctions and intersections has been proposed. This aid will help these people to detect, recognize and follow the signal with the help of machine learning algorithms. This will aid the destitute of vision to accurate perception of the

Fig. 6 Classification report of SVM-3 kernels
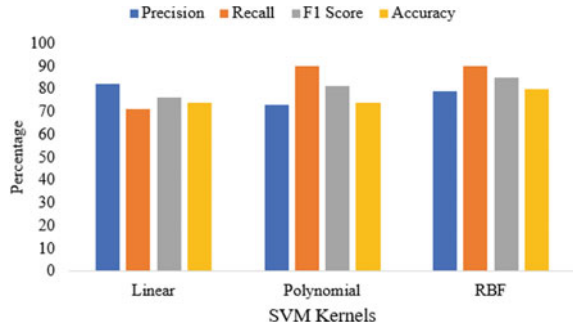
Fig. 7 Comparison of various classifiers
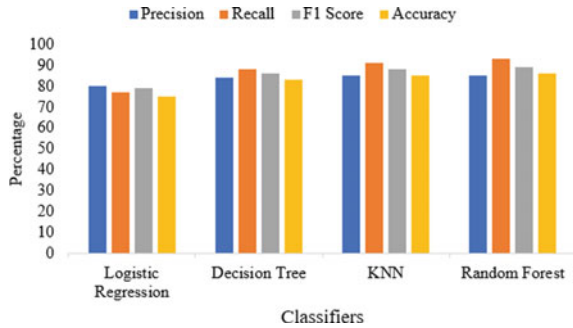
Table 3 Comparison of all models

	Precision	Recall	$F1$ score	Accuracy
Logistic regression	80	77	79	75
SVM RBF	79	90	85	80
KNN	85	91	88	85
Decision tree	84	88	86	83
Random forest	85	93	89	86

environment. The difference between the existing approaches and this approach is that pre-trained deep learning models have not been used. Instead, a model that runs solely on the computer vision techniques and feature-extraction based classification has been proposed.

The advantage of the proposed model is that the computational cost of training is lesser than the deep learning models. There is a tradeoff between accuracy and speed. However, accuracy is more crucial than speed in this application. The accuracy can be increased with further training and introducing variations in the dataset in future. Further, the model can also be made less computationally complex so that it can be deployed on embedded systems.

References

1. Statistics. https://www.orbis.org/
2. Ghilardi MC, Simões G, Wehrmann J, Manssour IH, Barros RC (2018) Real-time detection of pedestrian traffic lights for visually-impaired people. In: 2018 international joint conference on neural networks (IJCNN), pp 1–8.https://doi.org/10.1109/IJCNN.2018.8489516
3. Kaluwahandi S, Tadokoro Y (2001) Portable traveling support system using image processing for the visually impaired. J Instit Image Inf Televis Eng 55:337–340. https://doi.org/10.1109/ICIP.2001.959022
4. Chen Z, Huang X (2016) Accurate and reliable detection of traffic lights using multiclass learning and multi object tracking. IEEE Intell Transp Syst Mag 8(4):28–42. https://doi.org/10.1109/MITS.2016.2605381
5. Suda S, Ohnishi K, Iwazaki Y, Asami T (2018) Robustness of machine learning pedestrian signal detection applied to pedestrian guidance device for persons with visual impairment. In: 2018 12th France-Japan and 10th Europe-Asia congress on mechatronics, pp 116–121. https://doi.org/10.1109/MECATRONICS.2018.8495748
6. Wonghabut P, Kumphong J, Ung-arunyawee R, Leelapatra W, Satiennam T (2018) Traffic light color identification for automatic traffic light violation detection system. In: 2018 international conference on engineering, applied sciences, and technology (ICEAST), pp 1–4. https://doi.org/10.1109/ICEAST.2018.8434400
7. Lee S, Kim J, Lim Y, Lim J (2018) Traffic light detection and recognition based on Haar-like features. In: 2018 international conference on electronics, information, and communication (ICEIC), pp 1–4. https://doi.org/10.23919/ELINFOCOM.2018.8330598
8. Yang K, Cheng R, Bergasa LM, Romera E, Wang K, Long N (2018) Intersection perception through real-time semantic segmentation to assist navigation of visually impaired pedestrians. In: 2018 IEEE international conference on robotics and biomimetics (ROBIO), pp 1034–1039. https://doi.org/10.1109/ROBIO.2018.8665211
9. Ying J, Tian J, Lei L (2015) Traffic light detection based on similar shapes searching for visually impaired person. In: 2015 sixth international conference on intelligent control and information processing (ICICIP), pp 376–380.https://doi.org/10.1109/ICICIP.2015.7388200
10. Mascetti S, Ahmetovic D, Gerino A, Bernareggi C, Busso M, Rizzi A (2016) Robust traffic lights detection on mobile devices for pedestrians with visual impairment. Comput Vis Image Underst 148:123–135. ISSN: 1077-3142. https://doi.org/10.1016/j.cviu.2015.11.017
11. Omachi M, Omachi S (2009) Traffic light detection with color and edge information. In: 2009 2nd IEEE international conference on computer science and information technology, pp 284–287. https://doi.org/10.1109/ICCSIT.2009.5234518
12. da Silva Soares JC, Borchartt TB, de Paiva AC, de Almeida Neto A (2018) Methodology based on texture, color and shape features for traffic light detection and recognition. In: 2018 international joint conference on neural networks (IJCNN), pp 1–7. https://doi.org/10.1109/IJCNN.2018.8489669
13. Ozcelik Z, Tastimur C, Karakose M, Akin E (2017) A vision based traffic light detection and recognition approach for intelligent vehicles. In: 2017 international conference on computer science and engineering (UBMK), pp 424–429.https://doi.org/10.1109/UBMK.2017.8093430

14. Pongseesai C, Chamnongthai K (2019) Semantic traffic light understanding for visually impaired pedestrian. In: 2019 international symposium on intelligent signal processing and communication systems (ISPACS), pp 1–2. https://doi.org/10.1109/ISPACS48206.2019.8986355

15. Al-Nabulsi J, Mesleh A, Yunis A (2017) Traffic light detection for colorblind individuals. In: 2017 IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT), pp 1–6. https://doi.org/10.1109/AEECT.2017.8257737

16. Wu X-H, Hu R, Bao Y-Q (2018) Fast vision-based pedestrian traffic light detection. In: 2018 IEEE conference on multimedia information processing and retrieval (MIPR), pp 214–215. https://doi.org/10.1109/MIPR.2018.00050

17. Sooksatra S, Kondo T (2014) Red traffic light detection using fast radial symmetry transform. In: 2014 11th international conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON), pp 1–6. https://doi.org/10.1109/ECTICon.2014.6839767

18. Muslu G, Bolat B (2019) Nighttime vehicle tail light detection with rule based image processing. In: 2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT), pp 1–4. https://doi.org/10.1109/EBBT.2019.8741541

19. Binangkit JL, Widyantoro DH (2016) Increasing accuracy of traffic light color detection and recognition using machine learning. In: 2016 10th international conference on telecommunication systems services and applications (TSSA), pp 1–5. https://doi.org/10.1109/TSSA.2016.7871074

20. Mahendran JK, Barry DT, Nivedha AK, Bhandarkar SM (2021) Computer vision-based assistance system for the visually impaired using mobile edge artificial intelligence. In: 2021 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW), pp 2418–2427. https://doi.org/10.1109/CVPRW53098.2021.00274

21. Saleh S, Saleh H, Nazari M, Hardt W (2019) Outdoor navigation for visually impaired based on deep learning

22. Mallikarjuna GCP, Raju Hajare R, Pavan PSS (2021) Cognitive IoT system for visually impaired: machine learning approach. Mater Today: Proc. ISSN: 2214-7853. https://doi.org/10.1016/j.matpr.2021.03.666

23. Hsieh I-H, Cheng H-C, Ke H-H, Chen H-C, Wang W-J (2020) Outdoor walking guide for the visually-impaired people based on semantic segmentation and depth map. In: 2020 international conference on pervasive artificial intelligence (ICPAI), pp 144–147. https://doi.org/10.1109/ICPAI51961.2020.00034

24. Tapu R, Mocanu B, Zaharia T (2020) Wearable assistive devices for visually impaired: a state of the art survey. Pattern Recogn Lett 137:37–52. ISSN 0167-8655. https://doi.org/10.1016/j.patrec.2018.10.031

25. Ivanchenko V, Coughlan J, Shen H (2008) Crosswatch: a camera phone system for orienting visually impaired pedestrians at traffic intersections. Lect Notes Comput Sci 5105:1122–1128. https://doi.org/10.1007/978-3-540-70540-6_168

26. Tian Y (2014) RGB-D sensor-based computer vision assistive technology for visually impaired persons. https://doi.org/10.1007/978-3-319-08651-4_9

27. Pardasani A, Indi PN, Banerjee S, Kamal A, Garg V (2019) Smart assistive navigation devices for visually impaired people. In: 2019 IEEE 4th international conference on computer and communication systems (ICCCS), pp 725–729. https://doi.org/10.1109/CCOMS.2019.8821654

28. Kim S, Lee J, Ryu B, Lee C (2008) Implementation of the embedded system for visually-impaired people. In: 4th IEEE international symposium on electronic design, test and applications (delta 2008), pp 466–469. https://doi.org/10.1109/DELTA.2008.78

29. Dakopoulos D, Bourbakis NG (2010) Wearable obstacle avoidance electronic travel aids for blind: a survey. IEEE Trans Syst Man Cybern Part C (Appl Rev) 40(1):25–35

30. Froneman T, van den Heever D, Dellimore K (2017) Development of a wearable support system to aid the visually impaired in independent mobilization and navigation. In: 2017 39th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp 783–786. https://doi.org/10.1109/EMBC.2017.8036941

31. Shiizu Y, Hirahara Y, Yanashima K, Magatani K (2007) The development of a white cane which navigates the visually impaired. In: 2007 29th annual international conference of the IEEE Engineering in Medicine and Biology Society, pp 5005–5008. https://doi.org/10.1109/IEMBS.2007.4353464

32. Joe Louis Paul I, Sasirekha S, Mohanavalli S, Jayashree C, Moohana Priya P, Monika K (2019) Smart Eye for Visually Impaired—an aid to help the blind people. In: 2019 international conference on computational intelligence in data science (ICCIDS), pp 1–5. https://doi.org/10.1109/ICCIDS.2019.8862066

33. Dambhare S, Sakhare A (2011) Smart stick for blind: obstacle detection artificial vision and real-time assistance via GPS. In: 2nd national conference on information and communication technology (NCICT), pp 31–33

34. Rahman A, Nur Malia KF, Milan Mia M, Hasan Shuvo ASMM, Hasan Nahid M, Zayeem ATMM (2019) An efficient smart cane based navigation system for visually impaired people. In: 2019 international symposium on advanced electrical and communication technologies (ISAECT), pp 1–6. https://doi.org/10.1109/ISAECT47714.2019.9069737

35. Pedestrian Augmented Traffic Light Dataset. https://kaggle.com

36. LISA Traffic Light Dataset. https://kaggle.com

A Comparison of Different Machine Learning Techniques for Sentiment Analysis in Education Domain

Bhavana P. Bhagat and Sheetal S. Dhande-Dandge

Abstract For any educational institution, students are one of the most significant and important stakeholders. The role of teacher is to improve the intellectual level of the students. Teacher, student, and academic institution perform better when they use innovative teaching strategies. In the realm of education, the primary need of today is to monitor students understanding about learning and to improve or change teaching methodologies. Several research used various approaches to collect student's opinions in order to analyze feedback and present the results to teacher. To do this, we must investigate the application of various sentiment analysis techniques in order to determine student opinions from their comments. Thus, the purpose of this work is to investigate several approaches or mechanism for collecting students opinions in the form of feedback, via various platforms, analyzing and evaluating opinions using various machine learning and deep learning techniques. The paper also addresses the needs of a different stakeholders, like students, educators, and researchers, to improve teaching effectiveness by using sentiment analysis in education domain.

Keywords Sentiment analysis · Machine learning · Student's opinions · Education domain · Natural language processing (NLP)

1 Introduction

Teachers are the backbone of the education system. The performance of teachers is a central point of attention of foremost educational researchers. Teachers' effectiveness is judged not only by their qualifications and knowledge, but also by their style of teaching, dedication, commitment, and use of new tools to stay on the cutting edge in the classroom.

B. P. Bhagat (✉)
Department of Computer Engineering, Government Polytechnic, Yavatmal, MS, India
e-mail: bhavana.bhagat30475@gmail.com

S. S. Dhande-Dandge
Department of Computer Science & Engineering, Sipna College of Engineering and Technology, Amravati, MS, India

To address the changing demands of the classroom, effective teachers with a broad repertoire and the ability to employ various strategies skillfully are required. Knowing what students think of teaching is one of the most effective strategies for a teacher to improve teaching methodology. Thus, gaining access to students' opinions in the educational process entails allowing them to provide feedback on their teacher's performance, including their perspective on instruction, organization, classroom environment, and the amount learned quality.

One of the most essential strategies for assessing the quality of the educational process is teacher evaluation. It is mostly used in colleges and universities to assess teacher effectiveness and course delivery in higher education. An evaluation questionnaire can be used together with information. Quantitative data can be collected using closed-ended questions like MCQs, and qualitative data can be collected with open-ended questions like comments and suggestions from student's perspectives in textual form. Instructors frequently struggle to draw conclusions from such open-ended comments because they are usually loaded with observations and insight. Some of the student's responses may appear to be contradictory, for as when one student says one thing and another says the exact opposite. Quality education plays a very vital role nowadays in the growth of educational institutions. Also, the success of any educational institution depends on quality education, good academic performance, and retention of students. Students and their families do extensive online research by searching data to gain better knowledge of the prospective institution. Academic quality, campus placement, financial aid, campus facility, infrastructure, socialization, and educational policies are some of the key factors that students focus on before the admission process.

Thus, students are one of the most important key stakeholders for every educational institute. Enhancing the knowledge of student intellectual level is the responsibility of a teacher. Using innovative techniques in teaching improves the performance of students, teachers, and academic institutions. "The primary need of today is to monitor students understanding about learning and improving or changing teaching methodology in the field of the education domain. To ensure ongoing development in the teaching and learning experience, it is essential to assure that students' thoughts and feedback are taken seriously" [1].

Sentiment analysis is gaining popularity in various areas of text mining and natural language processing these days. Many industries, including education, consumer information, marketing, literature, applications, online review websites, and social media, have begun to analyze sentiments and opinions. Because of its widespread use, it has attracted the attention of various stakeholders, including customers, organizations, and governments, who want to examine and explore their perspectives. As a result, one of the most important functions of sentiment classification is to assess student feedback in the education area, as well as online documents such as blogs, comments, reviews, and new products as a whole, and categorize them as positive, negative, or neutral. The study of sentimental analysis has recently gained popularity among academics, and a number of research projects have been done.

2   Related Work

The education sector is undergoing a revolution to live up to standard in today's competitive world. With the generation of huge data, technologies have been developed to store and process data much easier. Thus, the use of sentiment analysis in the education domain will play a very important role for large users, namely teachers, students, and educational institutes. For analyzing the feedback, different machine learning languages and deep learning models will be used [1].

In this section, we are going to discuss existing work in the area of sentiment analysis in the education domain using machine learning and deep learning methods.

2.1   Methods

We started our survey by searching for relevant research studies on internet websites, i.e., digital libraries like IEEE Explore, Google Scholar, Research Gate, Springer, ACM Library, etc. The web search was conducted by using other search engines to trawl the digital libraries and databases. The key terms or search strings that we used were "Sentiment Analysis in education domain", "Effectiveness of teacher's performance using sentiment analysis", "sentiment analysis of student feedback", "sentiment analysis using machine learning". The above different terms are mainly used to search for research studies. Also, many articles were identified through scanning the reference list of each one of these articles. There is a lot of literature on the use of sentiment analysis in the education domain. Thus, studies published during the year 2014–2021 related to sentiment analysis in education sectors, tools, and techniques related to it were surveyed. Nearly 96 different papers have been referred, out of which nearly 80 papers were related to the use of sentiment analysis in different domains. Thus, 24 studies related to sentiment analysis using different methods specifically in the education domain have been included for further studies. Some of the important work in this area has been discussed as follow:

Some researchers [2–6] in their work used the lexicon-based approach for sentiment analysis. Many of them used labeled data. Thus, the procedure of the learning model is not required. But required powerful linguistic resources which are not always available. Accuracy performance is good as compared to other approaches. Some researchers [7–16] in their work used machine learning-based techniques. In this, the algorithm is first trained with some inputs with known outputs in order to work with new unknown data later. These algorithms work on large corpus, and accuracy performance is better for some algorithms, i.e., naïve Bayes and support vector machines as compared to other approaches. Some researchers [17–19], in their work, used hybrid approaches in order to improve accuracy performance. Finally, some researchers [20–23] used deep learning models as one of the dominating models in the education domain with better performance as compared to other models.

2.2   Recent Advances

The technique of extracting information about an entity and determining whether or not that entity has any subjectivities is known as sentiment analysis. The primary goal is to see if the user-generated text conveys positive, negative, or neutral emotions. The three stages of sentiment classification extraction are the document level, sentence level, and aspect level. To deal with the problem of sentiment analysis, there are three ways: lexicon-based techniques, machine learning-based techniques, and hybrid approaches. The original technique for sentiment analysis was lexicon-based, which was then divided into two methods: dictionary-based and corpus-based. Sentiment classification is performed using a dictionary of terms, such as those found in Senti-WordNet and WordNet, in the first method, whereas corpus-based sentiment analysis is based on the statistical analysis of the contents of documents, using different methods such as K-nearest neighbors, conditional random fields, and hidden Markov models, among others, in the second method. A summary of the salient features of some of the reviewed papers is shown in Table 1.

Traditional models and deep learning models are the two types of machine learning-based methodologies presented for sentiment analysis problems. Machine learning approaches such as the naïve Bayes, support vector machines, and maximum entropy classifiers are examples of traditional models. Lexical features, sentiment lexicon-based features, part of speech, and adjectives and adverbs are among the several inputs to these algorithms. The accuracy of systems is determined by which features are used as input. Deep learning models, when compared to regular models, can yield greater outcomes. Deep learning models that can be used for sentiment analysis include convolutional neural networks, deep neural networks, and recursive neural networks (RNN).

These methods identify flaws with classification at the document, sentence, or aspect level. The hybrid method combines methodologies based on lexical and machine learning. Sentiment lexicons are frequently used in the bulk of these systems [24]. Machine learning techniques have been employed for sentiment analysis in recent years to achieve classification either by supervised or unsupervised algorithms. Hybrid methods were also adopted on occasion. The majority of research use either machine learning or lexical techniques. Also, based on the above-mentioned reviews, it seems that neural networks are now dominating as a preferred method of most authors in their work in the education domain. Machine learning solutions have adopted deep network models such as long short-term memory, bidirectional LSTM, recursive neural networks, and convolutional neural networks.

3   Results and Discussion

From the above literature review, Table 2 shows the use of different approaches used for sentiment analysis. We can say that mostly sentiment analysis has been

Table 1 Summary of reviewed papers

S. No.	Year	Study	Method/algorithm	Data corpus	Performance
1	2014 Nov	Altrabsheh et al. [7]	Naïve Bayes, CNB SVM ME	1036 instances	F-score $= 0.94$, $P = 0.94$, $R = 0.94$, $A = 0.94$ (for SVM linear kernel unigram as compared to BI + TRI SVM, NB, CNB, ME, SVM Poly)
2	2014	Pong-Inwong and Rungworawut [8]	SVM, ID3, and naïve Bayes	175 instances from student opinions ab	Support vector machine accuracy—0.98, precision, recall, F-score—0.98
3	2016	Rajput et al. [2]	Lexicon-based method	1748 comments	$A = 91.2$, $P = 0.94$, $R = 0.97$, F-score $= 0.95$
4	2016	Balahadia et al. [9]	Naïve Bayes	Written comments	NP
5	2016	Sindhu et al. [20]	LSTM model of deep learning for layers 1 and 2	5000 comments	F-score $= 0.86$, $P = 0.88$, $R = 0.85$, $A = 0.93$
6	2017	Rani and Kumar [3]	NRC, emotion lexicon	4000 comments during course and 1700 after course	NP
7	Jun-17	Esparza et al. [10]	SVM linear, SVM radial, SVM polynomial	1040 comments in Spanish	B.A $= 0.81$, $A = 0.80$, sensitivity $= 0.74$ (for SVM linear), specificity $= 0.74$ (SVM radial)
8	2017	Sivakumar and Reddy [11]	Naive Bayes and K-means algorithm SentiWordNet	Online student comments	NP
9	2017	Nasim et al. [12]	Random forest and SVM	1230 comments	Proposed hybrid approach (TF-IDF with domain-specific lexicon) accuracy $= 0.93$, F-score $= 0.92$
10	2017	Aung and Myo [4]	Lexicon-based method	745 words	NP
11	2018	Gutiérrez et al. [13]	SVM linear, SVM radial, SVM polynomial	1040 comments in Spanish	SVM linear kernel-above 0.80
12	2018	Atif [5]	N-gram	Students responses	N-gram accuracy $= 0.80$

(continued)

Table 1 (continued)

S. No.	Year	Study	Method/algorithm	Data corpus	Performance
13	2018	Newman and Joyner [6]	VADER as SA tool	Students comments	NP
14	2018	Nguyen et al. [21]	LSTM and DT-LSTM models	UIT-VSFC corpus 16,175 sentences	LD-SVM $F1$-score 90.2% and accuracy 90.7%
15	2018	Cabada et al. [22]	CNN, LSTM	Yelp-147672 SentiText-10834 EduSere-4300 students comments	CNN + LSTM $A = 88.26\%$ in SentiText and $A = 90.30\%$ in EDuEras
16	2019	Lalata et al. [17]	Naïve Bayes, SVM DT RF	1822 comments	Best model NB with $A = 90.26$, SVM with $A = 90.20$, DT $A = 89$
17	2019	Chauhan et al. [14]	Naive Bayes classifier and online sentiment analyzer	1000 students comments	Performance using sentiment score F-score $= 0.72$, $P = 0.69$, $R = 0.76$ Performance teaching aspects F-score $= 0.80$, $P = 0.76$, $R = 0.84$
18	2020	Kaur et al. [15]	Naïve Bayes, SVM, emotion lexicon using Weka 3.7	4289 comments	SVM F-score $= 84.33$, NB F-score $= 85.53$
19	2020	Chandra and Jana [18]	CNN-RNN, LSTM	First GOP debate, bitcoin tweets, and IMDB movie reviews	For ML classifiers accuracy in range of 81.00–97, for CNN-RNN, LSTM, and their combination-accuracy range 85–97
20	2020	Sangeetha and Prabha [23]	LSTM, LSTM + ATT, multihead ATT and FUSION	16,175 students feedback sentences	FUSION model (multihead attention + embedding + LSTM) $A = 94.13$, $R = 88.72$, $P = 97.89$
21	2021	Qaiser [19]	Naïve Bayes, SVM DT DL	5000 comments	DL-A $= 96.41$ (best), NB-A $= 87.18$, SVM-A $= 82.05$, DT—68.21 (poor)

(continued)

Table 1 (continued)

S. No.	Year	Study	Method/algorithm	Data corpus	Performance
22	2021	Mabunda et al. [16]	Support vector machines, multinomial naive Bayes, random forests, K-nearest neighbors, and neural networks	Students feedback from Kaggle dataset 185 records	SVM—81% multinomial NB—81% random forest—81%, K-NN—78%, neural network—84%

Table 2 Sentiment analysis approaches used from 2014 to 2021

Methods/approaches	Work
Lexicon-based approach	[2–6]
Machine learning	[7–16]
Hybrid approaches	[18, 21, 23]
Deep learning	[20–23]

extensively applied in the education domain in the last five years. In the early years, lexicon-based techniques have been used for analyzing the sentiments of students. Later, different machine learning algorithms have been used for analyzing sentiment analysis. In recent years, the use of deep learning in the education domain is increased.

Due to the use of sentiment analysis in the education domain improved learning process in course performance, improved reduction in course retention improved teaching process, and satisfaction with the course of student increased. Also, according to the above literature presented, the most used techniques under the supervised machine learning approach are naive Bayes, support vector machine which provide higher precedence than other techniques discussed in the review.

After analyzing the use of different approaches, it is found that the accuracy performance of different approaches improved with the use of machine learning and deep learning approaches. Many researchers calculate the performance of their system using various performance measures, i.e., accuracy, precision, recall, and *F*-score. The accuracy performance of their models was mentioned by most of the researchers, from which we calculated the average accuracy performance for each of the different three methods and conclude that deep learning models are better as compared to lexicon-based and machine learning approaches. Further, recently the use of deep learning approaches such as long short-term memory (LSTM), bidirectional LSTM, recursive neural networks (RNN), and convolutional neural networks (CNN) provide higher precedence over machine learning approaches. Table 3 and Fig. 1 show the averaged accuracy of different methods/approaches used by researchers. Figure 1 depicts the fact that deep learning models have performed better than all other lexicon-based and machine learning classifiers used in this study.

Table 3 Averaged accuracy of different methods/approaches

S. No.	Methods/approaches	Average accuracy performance	No. of work
1	Lexicon-based	0.86	5
2	Machine learning	0.88	10
3	Deep learning	0.90	4

Fig. 1 Averaged accuracy of different methods
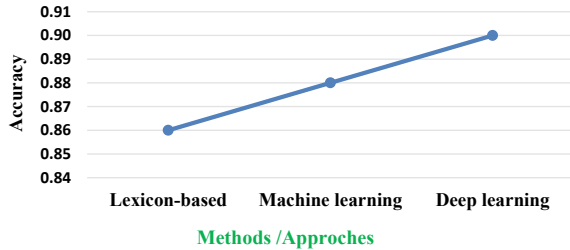
Moreover, from the above reviews, the dataset used in most of the studies is from online comments or from datasets available online on different repositories like Kaggle, etc.

Research Gaps: After reviewing previous work related to the education domain, some limitations were found as follows:

- Many of these works, such as Rajput et al. [2], Rani and Kumar [3], and Aung and Myo [4] used a lexicon-based approach that is less accurate than machine learning.
- Some of these researches, including Balahadia [9], Rani and Kumar [3], Shiv-Kumar et al. [11], Aung and Myo [4], and Newman et al. [6], were not evaluated using common sentiment analysis assessment metrics, such as accuracy, precision, recall, and F-score, and also they have not included any details of assessment or testing the performance of the model.
- Naive Bayes and support vector machine are mostly used techniques for the sentiment analysis.
- Also trends to combine machine learning and lexicon-based or machine learning and deep learning to perform better for the sentiment analysis process.
- However, still there is scope to explore for newer algorithms or optimized the existing ones for better accuracy.
- Some models do not provide visual results for sentiment analysis, nor do they provide any data or research to evaluate the visualization's usefulness and usability.
- The majority of research in this area focus on e-learning. However, there are distinctions between e-learning and sentiment analysis in the classroom.
- Very few models use primary dataset and use the machine and deep learning approaches along with evaluation metrics.

4 Conclusion

In this paper, we introduced the research done by various researchers related to the use of sentiment analysis in the educational domain to evaluate teacher's performance and to improve the teaching quality and overall performance of institutions. We investigated various approaches to sentiment analysis for different levels and situations. We also presented some work of deep learning in the educational domain for aspect-based sentiment analysis. Finally, we discussed some limitations related to previous work in the education domain.

References

1. Archana Rao PN, Baglodi K (2017) Role of sentiment analysis in education sector in the era of big data: a survey. Int J Latest Trends Eng Technol 022–024
2. Rajput Q, Haider S, Ghani S (2016) Lexicon-based sentiment analysis of teachers' evaluation. Appl Comput Intell SoftComput 2016
3. Rani S, Kumar P (2017) A sentiment analysis system to improve teaching and learning. Computer 50(5):36–43
4. Aung KZ, Myo NN (2017) Sentiment analysis of students' comment using lexicon based approach. In: 2017 IEEE/ACIS 16th international conference on computer and information science (ICIS). IEEE, pp 149–154
5. Atif M (2018) An enhanced framework for sentiment analysis of students' surveys: Arab Open University Business Program Courses Case Study. Bus Econ J 9(2018):337
6. Newman H, Joyner D (2018) Sentiment analysis of student evaluations of teaching. In: International conference on artificial intelligence in education. Springer, Cham, pp 246–250
7. Altrabsheh N, Cocea M, Fallahkhair S (2014) Sentiment analysis: towards a tool for analysing real-time students feedback. In: 2014 IEEE 26th international conference on tools with artificial intelligence. IEEE, pp 419–423
8. Pong-Inwong C, Rungworawut WS (2014) Teaching senti-lexicon for automated sentiment polarity definition in teaching evaluation. In: 2014 10th international conference on semantics, knowledge and grids. IEEE, pp 84–91
9. Balahadia FF, Fernando MCG, Juanatas IC (2016) Teacher's performance evaluation tool using opinion mining with sentiment analysis. In: 2016 IEEE region 10 symposium (TENSYMP). IEEE, pp 95–98
10. Esparza GG, de-Luna A, Zezzatti AO, Hernandez A, Ponce J, Álvarez M, Cossio E, de Jesus Nava J (2017) A sentiment analysis model to analyze students reviews of teacher performance using support vector machines. In: International symposium on distributed computing and artificial intelligence. Springer, Cham, pp 157–164
11. Sivakumar M, Reddy US (2017) Aspect based sentiment analysis of students opinion using machine learning techniques. In: 2017 international conference on inventive computing and informatics (ICICI). IEEE, pp 726–731
12. Nasim Z, Rajput Q, Haider S (2017) Sentiment analysis of student feedback using machine learning and lexicon based approaches. In: 2017 international conference on research and innovation in information systems (ICRIIS). IEEE, pp 1–6
13. Gutiérrez G, Ponce J, Ochoa A, Álvarez M (2018) Analyzing students reviews of teacher performance using support vector machines by a proposed model. In: International symposium on intelligent computing systems. Springer, Cham, pp 113–122
14. Chauhan GS, Agrawal P, Meena YK (2019) Aspect-based sentiment analysis of students' feedback to improve teaching–learning process. In: Information & communication technology for intelligent systems. Springer, Singapore, pp 259–266

15. Kaur W, Balakrishnan V, Singh B (2020) Improving teaching and learning experience in engineering education using sentiment analysis techniques. In: IOP Conference Series: Materials Science and Engineering, vol 834, No 1. IOP Publishing, p 012026
16. Mabunda JGK, Jadhav A, Ajoodha R (2021) Sentiment analysis of student textual feedback to improve teaching. In: Interdisciplinary research in technology & management. CRC Press, pp 643–651
17. Lalata JAP, Gerardo B, Medina R (2019) A sentiment analysis model for faculty comment evaluation using ensemble machine learning algorithms. In: Proceeding of 2019 international conference on big data engineering, pp 68–73
18. Chandra Y, Jana A (2020) Sentiment analysis using machine learning and deep learning. In: 2020 7th international conference on computing for sustainable global development (INDIACom). IEEE, pp 1–4
19. Qaiser S (2021) A comparison of machine learning techniques for sentiment analysis. Turk J Comput Math Educ (TURCOMAT) 12(3):1738–1744
20. Sindhu I, Daudpota SM, Badar K, Bakhtyar M, Baber J, Nurunnabi M (2019) Aspect-based opinion mining on student's feedback for faculty teaching performance evaluation. IEEE Access 7:108729–108741
21. Nguyen VD, Van Nguyen K, Nguyen NLT (2018) Variants of long short-term memory for sentiment analysis on Vietnamese students' feedback corpus. In: 2018 10th international conference on knowledge and systems engineering (KSE). IEEE, pp 306–311
22. Cabada RZ, Estrada MLB, Bustillos RO (2018) Mining of educational opinions with deep learning. J Univ Comput Sci 24(11):1604–1626
23. Sangeetha K, Prabha D (2021) Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM. J Ambient Intell Hum Comput 12(3):4117–4126
24. Dang NC, Moreno-García MN, De la Prieta F (2020) Sentiment analysis based on deep learning: a comparative study. Electronics 9(3):483

Design and Simulation of 2 × 2 Microstrip Patch Array Antenna for 5G Wireless Applications

Kasigari Prasad, A. Jeevan Reddy, B. Vasavi, Y. Suguna Kumari, and K. Subramanyam Raju

Abstract In this communication era, 4G is playing a significant role. Apart from this, the fifth-generation communication network is becoming increasingly apparent as it provides numerous applications. In this paper, the main focus is on the design of a 2 × 2 microstrip patch antenna array with inserted insets for the C-band applications and its analysis in terms of bandwidth, resonant frequency, VSWR, return loss, and other parameters. The proposed antenna is designed by using FR-4 epoxy substrate and simulated in HFSS software. The simulated results reveal increased bandwidth and return losses compared with 2 × 2 microstrip patch array antenna without insets. A detailed investigative analysis on a 2 × 2 microstrip patch array antenna with insets is performed and observed a return loss of − 23.27 dB and bandwidth of 150 MHz with a gain of 7.3 dB in this paper.

Keywords Array · Patch antenna · 5G · HFSS · Flame retardant-4 · VSWR · Return loss

1 Introduction

The demand for quick communication has resulted in the development of wireless communications. Wireless technology improves efficiency, coverage, and flexibility and reduces costs. Wireless communication is the transfer of data such as audio, video, images, and so on between two or more places that are not physically linked. Wireless operations allow services like long-distance communication which are not possible with cables [1]. Following the debut of the first-generation mobile network in the early 1980s, mobile wireless communication has evolved through numerous stages over the previous few decades. Because of the global desire for additional connections, mobile communication standards progressed quickly to accommodate too many users.

K. Prasad (✉) · A. Jeevan Reddy · B. Vasavi · Y. Suguna Kumari · K. Subramanyam Raju
Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: kpd@aitsrajampet.ac.in
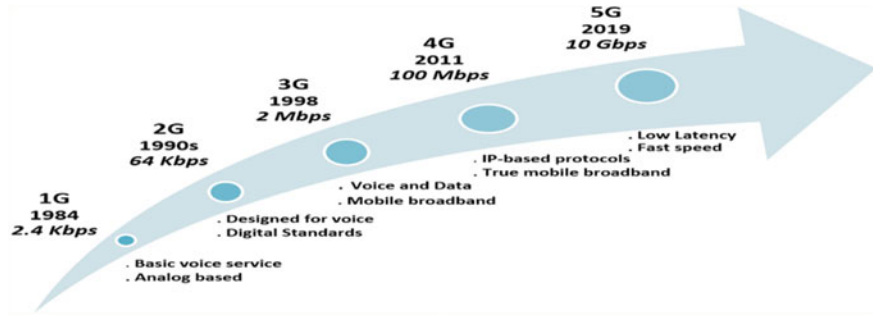
Fig. 1 Wireless communication technologies

Antennas are critical components of wireless communication. The information is delivered by the transmitter and received by the receiver via the antenna. An antenna is a type of transducer that emits or receives radio waves. Different varieties of antennas have been used in communication systems, with varying forms, sizes, orientations, and materials. Different types of antenna configurations are required for various purposes. Microstrip antennas are being used in communication equipment such as mobile phones and Wi-Fi devices [2]. 5G can provide faster data rates, greater resolution, reduced latency, more connection density, higher bandwidth, and higher power [3]. Figure 1 depicts the evolution of mobile communication.

2   Antenna Design

An antenna design is very critical in the investigation and application of the antenna in the specific communication system. To design an antenna that meets the required application, several antenna parameters are to be taken into consideration. The major antenna parameters are directivity, gain, bandwidth, radiation pattern, beam width, return loss, and VSWR. The other physical parameters like type of material, the height of material, shape, and size of the antenna including the cost of the antenna. In this way, the antenna is chosen carefully. Figure 2 depicts the antenna design procedure for the proposed antenna model.

Conduct a careful literature review on the topic of interest and identify the needed antenna parameters after a thorough examination of concepts. After that based on the selection, determine the antenna that best fits the set of specified parameters and design the suggested antenna within the constrained limits. The antenna is then theoretically modeled based on the set of specified parameters, and the suggested antenna is designed in an antenna modeling and simulation tool. The suggested antenna is developed and simulated in HFSS. Once the design has been finished in the HFSS, the antenna is simulated and the acquired results are compared with optimum

Fig. 2 Basic procedure for antenna design
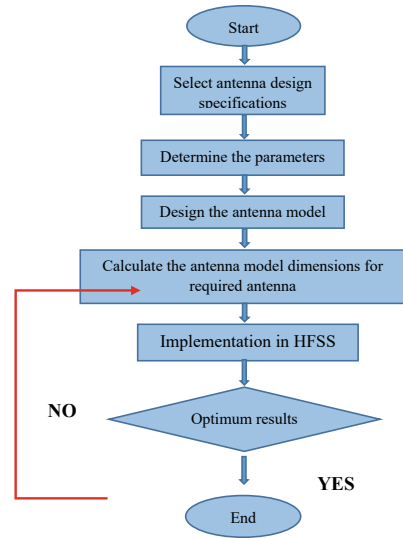
results. If the results are not met the requirements, then the antenna is remodeled to get optimum results. Finally, the collected results are subjected to results analysis.

3 Literature Overview

In the year 2021, Praveen Kumar Patidar and Nidhi Tiwari have studied microstrip patch antennas, and microstrip patch antenna arrays with various substrates theoretically and experimentally for 5G and investigated the frequency range of 5–60 GHz [4].

In the year 2021, Rafal Przesmycki et al., proposed a rectangular microstrip antenna for 5G applications in response to the growing demand for mobile data and mobile devices. The resonance frequency of this antenna is 28.00 GHz and reflectivity of 22.50 dB. The proposed antenna has a radiation efficiency of 80.18%, and the antenna gain for the resonance frequency is 5.06 dB. The results also show that its bandwidth is 5.57 GHz [5].

In the year 2020, Poonam Tiwari et al., proposed antenna that is simplistic and tiny, measuring $65 \times 65 \times 1.64$ (mm)^3. The intended antenna's tiny size allows it to be easily integrated into compact devices. The results show that the recurrence bandwidth spans the LTE band (4–7) GHz, with resonant frequencies of 4.91 GHz and 6.08 GHz, respectively, for VSWR less than 2, and S11 less than -10 dB [6].

In the year 2020, Manasa Buravalli et al., proposed a design, an antenna it has achieved a low return loss of -26.52 dB and a strong gain of 8.88 dB at the main lobes peak. Design is very small which is very compactable [7].

In the year 2020, Hangsa Raj Das et al., did a survey on frequently used approaches and designs which have been used by researchers for developing an efficient, tiny, compatible, and cost-effective microstrip antennas, which are primarily utilized in the building of reconfigurable, multiband, and wideband antennas, after which an initiator patch design is given with measurements on which technique will be adapted for the analysis of different antenna parameters [8].

In the year 2020, Rashmitha R. et al., presented a microstrip patch antenna capable of supporting 5G. At 43.7 GHz, this antenna operates in the extremely high-frequency spectrum. The small-sized antenna may be used in communication devices as well as smaller base stations [9].

In the year 2019, Ranjan Mishra et al., proposed a microstrip square-shaped antenna with stub feedline, 500 MHz of good bandwidth, and − 24 dB of high return loss [10].

In the year 2018, Wen-Shan Chen and Yung-Chi Lin proposed an antenna that is simple and compact in size. This antenna resonates at 3.45 and 3.57 GHz. The maximum gain is 5.37 dBi achieved about 30% of radiation efficiency. This design is suitable for 5G applications [11].

In the year 2018, Prasad K. et al., presented a blade antenna that works at frequencies spanning from 0.5 to 2 GHz. The examination of several blade antenna types and the determination of the appropriate antenna properties for airborne applications [12].

In the year 2016, K. Prasad et al., proposed a design it determined that when the length of the microstrip line feed grows, the antenna's characteristics are altered. As the antenna's frequency is increased, so are its directivity and gain. Return loss diminishes as frequency rises. The suggested antenna is suited for WIFI (2.4–2.48 GHz), wireless communication, and other applications [13].

In the year 2016, Prasad K. et al., presented a paper that provides an in-depth look of the blade antenna, a relatively new concept in airborne applications. Because of their small size, lighter weight, and aerodynamic design, blade antennas are ideal for airborne applications. The many types of blade antennas, their uses, benefits [14, 15].

4   2 × 2 Microstrip Patch Array Antenna Without Insets

The potential design for existing work is shown in Fig. 3, which is also a 2 × 2 antenna array arrangement without insets. The design specifications for the existing work are shown in Table 1. This antenna is resonated at frequencies of 3.45 and 3.57 GHz with gain of 5.37 dB, having a return loss of − 15 dB. Figure 4 depicts the results of this antenna model.
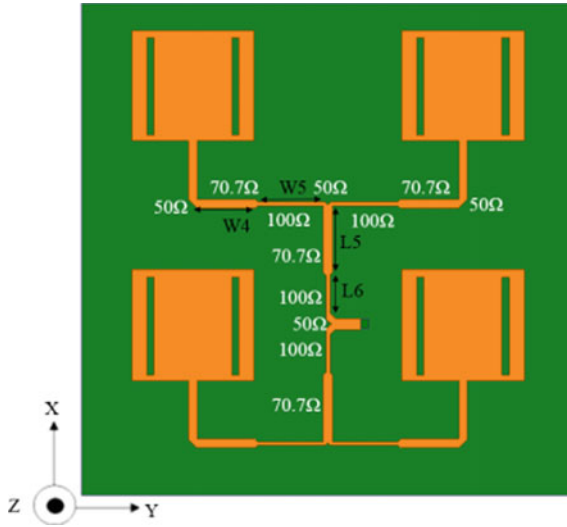
Fig. 3 Previous work design

Table 1 Design parameters for existing antenna

S. No.	Parameter	Value
1	Low frequency (f_l)	3 GHz
2	High frequency (f_h)	4 GHz
3	Dielectric constant (ε_r)	4.4/FR4-epoxy
4	Ground ($L \times W$)	88.5 mm × 88.5 mm
5	Substrate ($L \times W$)	88.5 mm × 88.5 mm
6	Substrate height (h)	1.6 mm
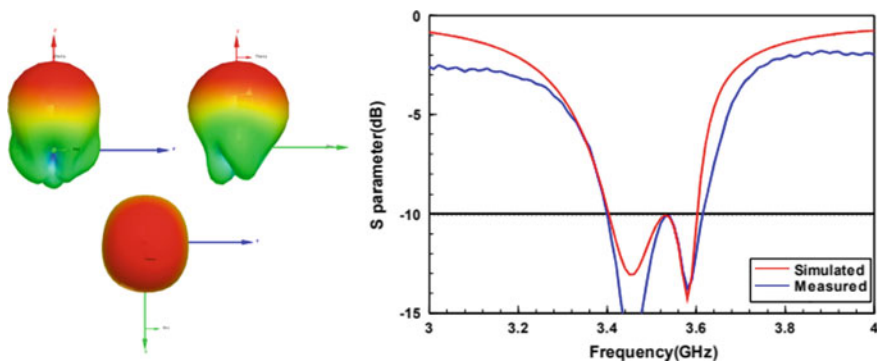
Fig. 4 Radiation pattern and S parameter results

5 2 × 2 Microstrip Patch Array Antenna with Insets

The view from the top of a proposed array microstrip patch antenna with insets is shown in Fig. 4, with one side of a dielectric substrate acting as a radiating patch and the other as a ground plane. Together, the patch and ground plane generate fringing fields, which are responsible for the antenna's emission. Due to the compact size, we recommended a 2 × 2 antenna array design. When the array size grows, the overall antenna size rises as well.

The geometry of the suggested design of a 2 × 2 microstrip patch array for C-band applications is seen in Fig. 5. The design is 88.5 mm × 88.5 mm × 1.6 mm (L W H) and is printed on flame retardant-4 having a 4.4 relative permittivity. Table 2 displays the dimensions for the array antenna. The antenna array for 5G C-band applications is made of rectangular microstrip with two slots and insets.
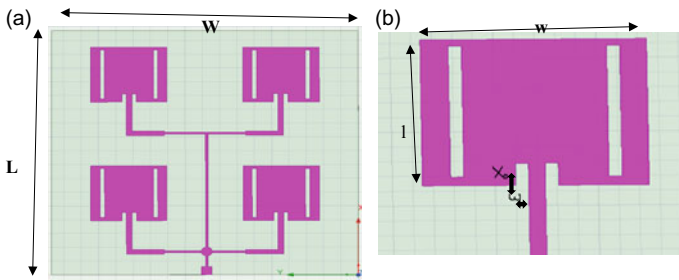
Fig. 5 **a** Top view. **b** Basic design of patch antenna

Table 2 Proposed antenna design parameters

S. No.	Parameter	Value
1	Low frequency (f_l)	3 GHz
2	High frequency (f_h)	6 GHz
3	Dielectric constant (ε_r)	4.4/FR4
4	Ground ($L \times W$)	88.5 mm × 88.5 mm
5	Substrate ($L \times W$)	88.5 mm × 88.5 mm
6	Substrate height (h)	1.6 mm
7	Single patch ($l \times w$)	19.8 mm × 22 mm
8	Inset ($X_\mathrm{o} \times \mathbf{w}$)	3 mm × 1.3 mm
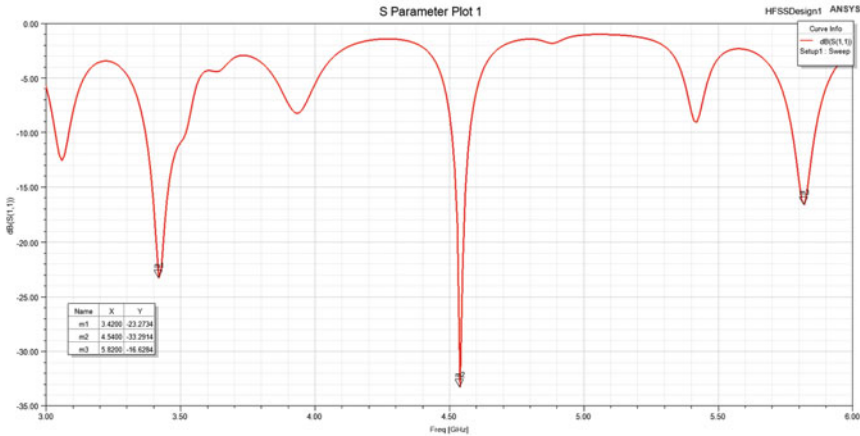
Fig. 6 Return loss

6 Results

6.1 S Parameter

S parameter (return loss) indicates the amount of power reflected by the antenna so it is known as the reflection coefficient or return loss. If S11 is 0 dB, then their will be no radiation from the antenna, all power is reflected back from the antenna. The proposed antenna has a return loss of − 23.27 dB at 3.42 GHz, − 33.29 dB at 4.5 GHz, and − 16.2 dB at 5.82 GHz. Figure 6 depicts the return loss.

6.2 Bandwidth

The bandwidth is defined as "the range of frequencies throughout which the antenna may send or receive information in the form of EM waves while correctly operating." The bandwidth of a broadband antenna is sometimes described as the ratio of allowable upper-to-lower operating frequencies. The bandwidth of a broadband antenna is stated as a percentage of the frequency difference across the center frequency bandwidth. The proposed antenna has a 150 MHz (3.52–3.37 GHz) bandwidth for the first band, the second band of 70 MHz (4.58–4.51 GHz), and 90 MHz (5.86–5.77 GHz) for the third band as shown in Fig. 7.
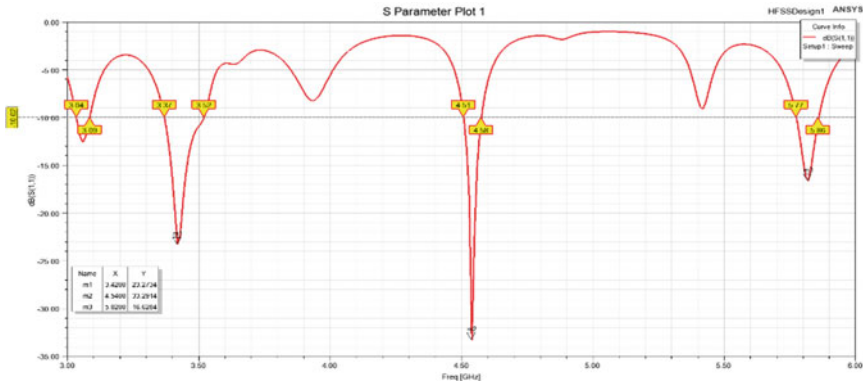
Fig. 7 Bandwidth calculation

6.3 Voltage Standing Wave Ratio (VSWR)

VSWR is a measurement of the degree of mismatch between a feedline and an antenna. When mounting and configuring transmitting antennas, measuring VSWR is the most prevalent use. When a feed line links a transmitter to an antenna, the impedance of the feed line and the antenna must be precisely matched for optimal energy transfer from the feed line to the antenna. Standing wave patterns are caused by the interaction of advancing waves with reflected waves. The range of VSWR values ranges from 1 to infinity (Fig. 8). A VSWR of less than 2 is found to be more appropriate for the majority of antenna applications.
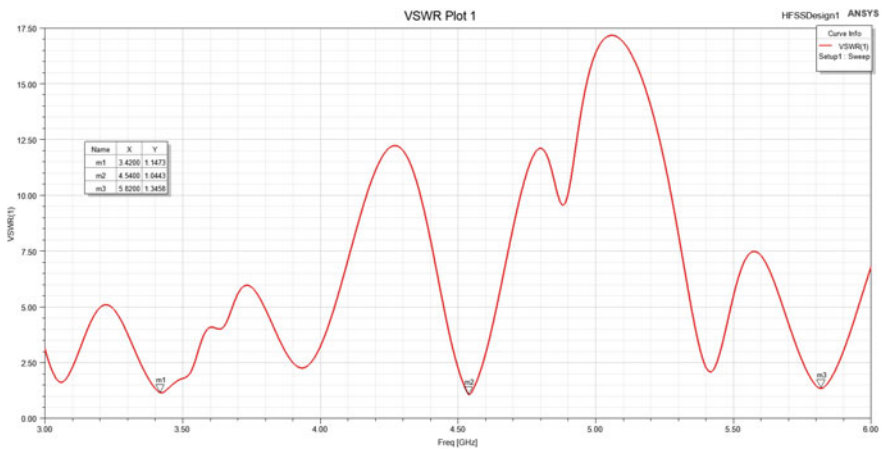
Fig. 8 Voltage standing wave ratios

Fig. 9. 3D radiation plot

Main lobe magnitude= 25.1dB

rE(theta)

rE(phi)
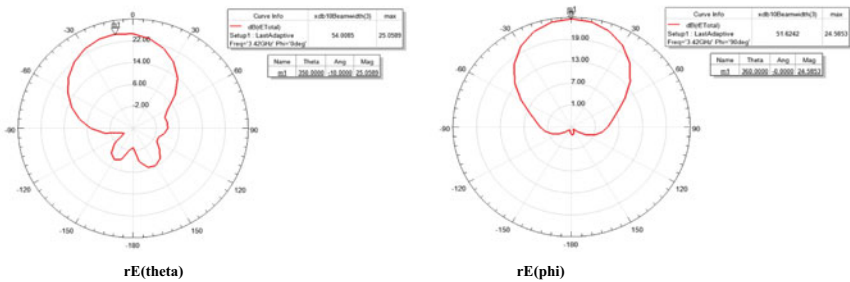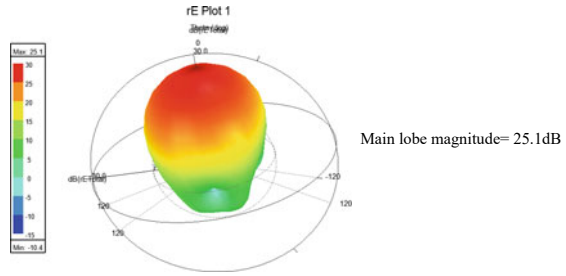
Fig. 10 Radiation plots of Theta and Phi planes

6.4 Radiation Plots

Radiation pattern of an antenna represents the energy emitted by the antenna. Radiation patterns are graphical form representations of how radiated energy is distributed in space as a function of direction. Figure 9 depicts the 3D radiation plot. Figure 10 depicts radiation plots of the Theta and Phi planes.

6.5 Gain

The gain of an antenna indicates how successfully it transforms input power into radio waves that go in a certain direction. The simulated gain maximum value is 7.3 dB. Figure 11 depicts the gain plot.

Fig. 11 Gain plot

Maximum Gain = 7.3dB
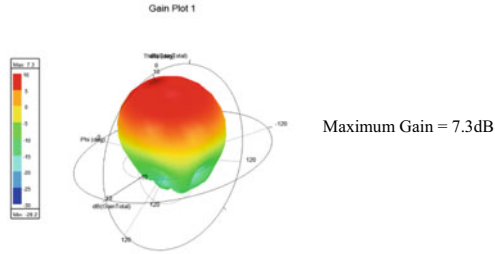
6.6 Directivity

Directivity is "the ratio of the subject antenna's highest radiation intensity to that of an isotropic or reference antenna emitting the same total power." This antenna has a directivity of 10.9 dB as shown in Fig. 12.

Table 3 compares proposed antenna findings to earlier design results in terms of bandwidth, return loss, resonant frequency, VSWR, etc.

Fig. 12 Directivity plot

Maximum Directivity = 10.9 dB
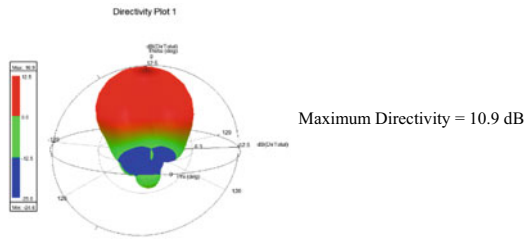
Table 3 Comparison of results of the previous design with the proposed design

Parameter	2 × 2 microstrip patch array antenna without insets	2 × 2 microstrip patch array antenna with insets
Antenna array	2 × 2	2 × 2
Bandwidth	80 MHz	150, 70, and 90 MHz
Return loss	− 15 dB	− 23.27, − 33.29, − 16.62 dB
Resonant frequency	3.45 and 3.57 GHz	3.42, 4.54 and 5.82 GHz
Gain	5.37 dB	7.3 dB
VSWR	> 1	1.14
No of band	Multi	Multi
Application	Wireless communication	Wireless communication

7 Conclusion

A rectangular microstrip patch antenna is designed and simulated using HFSS simulation software. The findings of the simulation are presented and analyzed. The proposed antenna structure is simple and small in size of 88.5mmx88.5mmx1.6mm. The intended antenna's tiny size makes it easy to utilize in small devices. Results show that the recurrence bandwidth covers the C-band, at resonant frequencies 3.42, 4.54, and 5.82 GHz individually for VSWR under 2. The antenna array's simulated and measured S11 satisfies the 5G C-band spectrum. In the above working band, it shows good impedance coordinating and radiation patterns. Hence, the proposed antenna can be used for remote correspondence applications in the C-band. The intended antenna performs well, with increased return loss and bandwidth.

References

1. Rappaport T (1996) Wireless communications: principles and practice, 2nd edn. Pearson, Upper Saddle River, NJ
2. Grieg DD, Engelmann HF (1952) Microstrip: a new transmission technique for the kilomegacycle range. Proc IRE 40(12):1644–1650
3. Warren D, Dewar C (2014) Understanding 5G: perspectives on future technological advancements in mobile. GSMA Intell 1–26
4. Patidar PK, Tiwari N (2021) A review paper on microstrip patch antenna (MPA) for 5G wireless technology. Turk. J. Comput. Math. Educ. 12(13):1741–1747
5. Przesmycki R, Bugaj M, Nowosielski L (2021) Broadband microstrip antenna for 5G wireless systems operating at 28 GHz. Electronics 10:1
6. Tiwari P, Gupta A, Nagaria D (2020) Design and analysis of 2×2 micro strip patch antenna array for 5G C-band wireless communication. In: 2020 JETIR Feb 2020, vol 7
7. Buravalli M et al (2020) Simulation study of 2×3 microstrip patch antenna array for 5G applications. IEEE
8. Das HR, Dey R, Bhattacharya (2020) A review paper on design for microstrip patch antenna. Topics in Intelligent Computing and Industry Design (ICID) 2(2); (2020) 166–168
9. Rashmitha R, Niran N, Abhinandan Ajit Jugale, Mohammed Riyaz Ahmed, Microstrip patch antenna design for fixed mobile and satellite 5G communications. Procedia Computer Science 171 (2020) 2073–2079
10. Mishra R, Mishra RG, Chaurasia RK, Shrivastava AK (2019) Design and analysis of microstrip patch antenna for wireless communication. Int J Innov Technol Expl Eng (IJITEE) ISSN: 2278-3075 8(7)
11. Chen W, Lin Y (2018) Design of 2x2 microstrip patch array antenna for 5G C-band access point applications. IEEE
12. Prasad K, Singh RP, Sreekanth N (2018) Design and analysis of blade antenna using diverse rectangular substrates with SSRR And SRR. Int J Pure Appl Math 118(14)
13. Prasad K, Lakshmi Devi B (2016) Design and simulation of wide-band square microstrip patch antenna. IJARCCE 5(7)
14. Prasad K, R.P.Singh, N.Sreekanth, "A Systematic survey on Blade antennas for Airborne applications", IJATEST, Volume.4, Special Issue.1Dec.2016
15. Prasad K, Singh RP, Sreekanth N (2017) Modeling and investigation of blade antenna with and without single ring resonator. IJARSE 6(1)

Fake Currency Detection: A Survey on Different Methodologies Using Machine Learning Techniques

Swathi Mattaparthi, Sheo Kumar, and Mrutyunjaya S. Yalawar

Abstract In the present world, everyone seems to be specializing in technologies; however, they're not worrying concerning security. Most users and researchers are giving more attention of the machine learning applications and technology itself, however less attention toward genuinely of the detection of currency notes. In this paper, we represent overview of the various methods and the techniques of various fake currency detection using machine learning techniques. This paper for the most part centers around the security examination of the innovation and uses of the machine learning algorithms, and condenses the vulnerabilities and anticipates potential assaults and furthermore outlines the present status of detection of fake currency in terms of security protection.

Keywords Fake currency · CNN · SVM · Data security · Privacy protection

1 Introduction

In our country, three times demonetization was done in years 1949, 1996, and recently in the year 2016. Demonetization is the situation where the government legally bans the notes and coins of a certain denomination. Nowadays we know that technology is a part of everyone life. Sometimes same technology brings some negative effects on individuals' life. Among them one of the negative impacts is counterfeit currency causes inflation. This is the common problem facing by every country. This shows effect on economy and reduces the value of real note. This leads to currency devaluation. Mostly, it will affect poor people and uneducated people in their regular times. Therefore, RBI using a new technology after every demonetization and adding some new features to currency notes. Watermarking, latent picture, microlettering, see through register optically variable connection, security threads, intaglio printing, fluorescence, identification mark. Until now, many techniques have been proposed to identify the currency note. There is a need to build a reliable system that can identify

S. Mattaparthi (✉) · S. Kumar · M. S. Yalawar
CMREC, Hyderabad, India
e-mail: swathilaxman0515@gmail.com

currency with minimal processing time and accuracy. We apply here a simple algorithm which works properly and gives more accuracy. Convolution neural network (CNN) plays a vital role in the recognition process and can improve the accuracy of the overall training through using CNN model. The move toward consists of works including image processing, edge detection, image segmentation and characteristic extraction and comparing images. The preferred outcome will be text output of the notes recognized and confirmed. These types of approaches are detection efficiency less and time-consuming too.

2 Literature Review

We will provide a literature overview of the few papers we examined for the subject as follows:

The paper titled as "Indian Currency Denomination Recognition and Fake Currency Identification" (2021) presented by B. Padmaja, P. Bhargav Naga Shyam, H. Ganga Sagar, B. Diwakar Nayak, and M. Bhushan Rao [1] makes use of three-layer CNN model with rectified linear unit (ReLU) activation function and two max pooling layers with OpenCV, implemented in Python.

The paper [2] titled as "Fake Currency Detection Application" (2021) presented by Aakash Vidhate, Yash Shah, Ram Biyani, Himanshu Keshri, and Rupali Nikhare and the paper titled as "Fake Indian Currency Recognization" (2021) presented by Ms. Zarina Begam, Dr. S. A. Quadri, Mr. Sayed Md Abrar Qazi [3] make use of notes detection system for recognition and extraction the features of banknote.

The paper [4] titled as "Detection of Fake Currency using Image Processing" (2019) presented by Ankush Singh, Prof. Ketaki Bhoyar, Ankur Pandey, Prashant Mankani, and Aman Tekriwal makes use of SVM Algorithm and using cloud storage for execution of image processing.

The paper [5] titled as "Analysis of Banknote Authentication System using Machine Learning Techniques" (2018) presented by Sumeet Shahani, Aysha Jagiasi, and Priya R. L. makes use of supervised machine learning techniques such as BPN and SVM. A recognition system should be installed to detect legitimacy of the note.

The paper [6] titled as "Research on Fake Indian Currency Note Detection using Image Processing" (2021) presented by Miss. I. Santhiya Irulappasamy makes use of algorithms Structure Similarity Index Metric (SSIM), Adaptive histogram equalization, Fast Discrete Wavelet Transform (FDWT), and Gray Level Co-occurrence Matrix (GLCM).

The paper [7] titled as "Bank Note Authentication and Classification Using Advanced Machine Algorithms" (2021) presented by Prof. Alpana D. Sonje, Prof. Dr. Harsha Patil, Prof Dr. D. M. Sonje makes use of Multi-Layer Perceptron Neural Network (MLPNN), Naïve Bayes, and Random Forest (RF) Algorithms on the standard data set.

The paper [8] titled as "Fake Currency Detection Using Image Processing" (2021) presented by Ankur Saxena, Pawan Kumar Singh, Ganesh Prasad Pal, Ravi Kumar

Tewari and the paper [9] titled as "Indian Counterfeit Banknote Detection using Support Vector Machine" presented by Santosh Gopane and Radhika Kotehca make uses of Machine Learning technique Support Vector Machine to authenticate the banknote.

The paper [10] titled as "Fake currency detection: A Survey" (2020) presented by Arun Anoop M, Dr. K. E. Kannammal, the paper [11] titled as "Review on Detection of Fake Currency using Image processing Techniques" (2021) presented by Dr. S. V. Viraktamath, Kshama Tallur, Rohan Bhadavankar, Vidya, and the paper [12] titled as "Fake Indian Currency Recognition System by using MATLAB" (2019) presented by Mr. S. S. Veling, Miss. Janhavi P. Sawal, Miss. Siddhi A. Bandekar, Mr. Tejas C. Patil, Mr. Aniket L. Sawant make use of techniques like counterfeit detection pen and MATLAB.

The paper [13] titled as "Banknotes Counterfeit Detection Using Deep Transfer Learning Approach" (2020) presented by Azra Yildiz, Ali Abd Almisreb, Šejla Dzakmic, Nooritawati Md Tahir, Sherzod Turaev, Mohammed A. Saleh makes use of deep learning technique in detecting the counterfeited BAM banknotes utilizing CNN models.

The paper [14] titled as "A Novel Approach for Detection of Counterfeit Indian Currency Notes Using Deep Convolutional Neural Network" (2020) presented by S. Naresh Kumar, Gaurav Singal, Shwetha Sirikonda, R. Nethravathi makes use of three-layered CNN-based model.

The paper [15] titled as "The Detection of Counterfeit Banknotes Using Ensemble Learning Techniques of AdaBoost and Voting" (2020) presented by Rihab Salah Khairy, Ameer Saleh Hussein, and Haider TH. Salim ALRikabi makes use of AdaBoost, and voting ensemble is deployed in combination with machine learning algorithms.

3 Discussion of Drawbacks

After studying all the reviewed articles, we came to understand the following points.

First of all, most of the articles chose to use CNN to recognize currency notes after comparing with other previous works because it can get more accuracy.

The use of CNN has several advantages, including, such as CNN is well-known for its architecture, and the best part is that no feature extraction is needed. The main advantage of CNN over its predecessor is that it can identify crucial feature without requiring people interaction.

Some drawbacks of CNN are: images in different positions are classified; because of operations like max pool, a convolutional neural network is substantially slower. If the CNN has many layers, the training phase can take a long time; if the machine does not have a powerful GPU and a CNN needs a big dataset to process and train the neural network.

A brief survey of the studied papers is presented in Table 1.

Table 1 Survey of papers

Paper title	Algorithms used	Drawbacks
1	CNN model with rectified linear unit (ReLU) activation function and two max-pooling layers	In the future, authors may use flip side features to improve their results
2	Machine learning algorithms for image detection and processing (CNN)	Authors want to implement the system for foreign currencies and tracking of device's location by including a module for currency conversion
3	Authors used two parts: currency recognition using image acquisition, preprocessing, edge detection, picture segmentation, and currency verification using Haar skin extraction and feature assessment	Authors will work with new classifier for their next work
4	Support vector network and used cloud storage for execution	Authors planning to proposed system could replace the hardware system in some initial stages of currency verification process
5	Machine learning techniques like backpropagation neural network and support vector machine using kernel functions	According to authors analysis BPN gives 100% detection rate than SVM
6	Using algorithms structure similarity index metric (SSIM), fast discrete wavelet transform, gray level co-occurrence matrix (GLCM), and artificial neural network, those are used to detect currency value	Authors may use different methods to detect their research in the future
7	MLPNN, Naïve Bayes algorithm, random forest algorithm using WEKA environment	In the future, the performance of the proposed algorithms can be improved by using the combination of various algorithms
8	KNN classifier using MATLAB	MATLAB specializes in matrix operations which is very important for researchers to modify its features for its high-performance measures
9	The proposed approach follows an image processing technique followed by the machine learning technique and with the use of support vector machine	Use of deep learning techniques with large amount of training data may be applied for better predictions
10	Machine learning algorithms	Author wants to design a new automatic system based on deep convolution neural network

Table 1 (continued)

Paper title	Algorithms used	Drawbacks
11	Calculation of mean intensity of RGB channels, Using Unified Modeling Language (UML) and HSV image, enhancement of Sift algorithm, KNN technique, ORB and BF matcher in OpenCV superresolution method, K-means algorithm and SVM algorithm	Author wants to increase the currency denomination and images should be taken from both sides and different angles to increase accuracy, and it should be cost-efficient and takes less time

4 Conclusion

In this paper, we focused after reviewing the above articles, we understood that deep learning techniques are crucial to currency note recognition. Based on deep learning literature review, most of the researchers have used convolutional neural networks for recognizing fake notes from the real notes. This technique can get more accuracy after comparing this technique with other previous works. Future researchers should build a system that can be used for currency recognition for all countries around the world because some of the models identify few different currencies. On the other hand, some of the current models cannot identify fake money. Furthermore, the researchers should create an application model for cellphones and web application to recognize fake and real money, especially for people with visual disabilities. Finally, the researchers should crucially work on extracting security thread features.

References

1. Indian currency denomination recognition and fake currency identification (2021) AMSE 2021. J Phys: Conf Ser. https://doi.org/10.1088/1742-6596/2089/1/012008
2. Fake currency detection application (2021) Int Res J Eng Technol (IRJET) 08(05). www.irjet.net
3. Fake Indian currency recognization (2021) In: International conference on artificial intelligence and machine learning, vol 8, no 5. ISSN: 2456-3307. www.ijsrcseit.com
4. Detection of fake currency using image processing (2019) Int J Eng Res Technol (IJERT) 8(12). ISSN: 2278-0181
5. Analysis of banknote authentication system using machine learning techniques (2018) Int J Comput Appl (0975-8887) 179(20)
6. Research on fake Indian currency note detection using image processing (2021) IJSDR 6(3). ISSN: 2455-2631
7. Bank note authentication and classification using advanced machine algorithms (2021) J Sci Technol 06(01). ISSN: 2456-5660
8. Fake currency detection using image processing (2018) Int J Eng Technol 7(4.39):199–205. www.sciencepubco.com/index.php/IJET
9. Indian counterfeit banknote detection using support vector machine. Available at: https://ssrn.com/abstract=3568724
10. Fake currency detection: a survey. Gedrag Organisatie Rev 33(04). ISSN: 0921-5077. http://lemma-tijdschriften.com/

11. Review on detection of fake currency using image processing techniques (2021) In: Proceedings of the fifth international conference on intelligent computing and control systems (ICICCS 2021). IEEE Xplore Part Number: CFP21K74-ART. ISBN: 978-0-7381-1327-2
12. Fake Indian currency recognition system by using MATLAB (2019) Int J Res Appl Sci Eng Technol (IJRASET) 7(IV). ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887. Available at: www.ijraset.com
13. Banknotes counterfeit detection using deep transfer learning approach. Int J Adv Trends Comput Sci Eng. Available at: http://www.warse.org/IJATCSE/static/pdf/file/ijatcse17295 2020.pdf
14. A novel approach for detection of counterfeit Indian currency notes using deep convolutional neural network (2020) IOP Conf Ser: Mater Sci Eng 981:022018. https://doi.org/10.1088/1757-899X/981/2/022018
15. The detection of counterfeit banknotes using ensemble learning techniques of AdaBoost and voting (2021) Int J Intell Eng Syst 14(1). https://doi.org/10.22266/ijies2021.0228.31

Bi-directional DC-DC Converters and Energy Storage Systems of DVR—An Analysis

A. Anitha and K. C. R. Nisha

Abstract Ensuring the quality of power supply has become a challenging task due to the intermittent nature of solar PV and wind turbine based power generation systems. Dynamic voltage restorer (DVR) is one of the cost-effective solutions to overcome most of the power quality (PQ) issues. DVR with energy storage topology suits ideally for deep voltage sags but results in increased complexity, converter rating and overall cost. Use of energy storage devices and bi-directional DC-DC converter helps to deliver quality power to consumers. Bi-directional topologies occupy lesser system space and deliver increased efficiency and better performance. In this paper, DVR topologies, different energy storage elements and power converters used in DVR are analyzed and reported.

Keywords Bi-directional DC-DC converter · Energy storage · DVR · UCAP · SEMS · Fuel cell

1 Introduction

The renewable sources such as wind and solar gaining popularity in the power generation arena results in the degradation of Grid's power quality due to their inherent nonlinearity. Voltage sag, swell and harmonic disturbances resulting by unexpected short-circuit failures or non-linear load fluctuations are the most common voltage quality disturbances. They may end in the termination of a process or the loss of data in digital equipment, resulting in further loss to power consumers.

Custom power devices have evolved to be more effective in managing power quality issues. DVR is a custom power device which proves to be cost-effective in overcoming voltage based PQ problems [1]. DVRs eliminate most of the sags and reduce the chances of load tripping caused by deeper sags, but they have a number of disadvantages, including standby losses, equipment costs, and the safety strategy necessary for downstream short circuits.

A. Anitha (✉) · K. C. R. Nisha
New Horizon College of Engineering, Bangalore, India
e-mail: anithanelson24@gmail.com

Different system topologies of DVR are realized including energy storage and with no energy storage element [2]. Energy storage design ranks second best in terms of performance, especially for severe voltage sags, but it has substantial downsides in terms of converter rating, complexity, and total cost (storage element and power converter).

If converter rating and energy storage elements are selected properly then DVR with energy storage topology can be ranked first in all aspects. The use of energy storage devices such as ultracapacitor, fuel cell stabilizes grid variations, and allowing users to get consistent electricity. Plug-in hybrid electric vehicle (PHEV) charging stations and Smart grid use the bi-directional DC to DC converter nowadays [3]. As a result, low-cost, efficient, and reliable bi-directional DC-DC converters and energy storage element are critical in today's environment.

This paper reviews topologies, energy storage system and bi-directional converters used in DVR. The following is the structure of this paper: Sect. 2 presents the different DVR topologies and its merits and demerits. Section 3 describes DVR with energy storage elements, Sect. 4 describes the comparison of storage elements and converter, and Sect. 5 compares DVR with energy and power converter.

2   DVR Topologies

DVR system topologies are classified into two main groups: topologies with no energy storage and topologies including energy storage [2]. The benefits and drawbacks of the topologies are discussed in Table 1.

The load side connected converter has been ranked first from the analysis [2] but the grid effects and rating of the series converter are higher. Hence not suitable in grid stabilization. The energy storage design ranks second best in terms of performance, especially for severe voltage sags, but it has substantial downsides in terms of rating of converter and storage systems. If the drawbacks are met then energy storage topology holds good for power quality applications.

3   DVR with Energy Storage Element

DVR with energy storage topology include energy storage element, inverter and injection transformer as shown in Fig. 1.

During sag, the inverter connected in series with the grid draws power from energy storage element and injects the missing voltage. Batteries energy storage system (BESS), superconducting magnetic energy storage (SMES), flywheels energy storage system (FESS), ultra capacitors (UCAPs), and fuel cell [4] are all viable rechargeable storage options used for integration into DVR to mitigate the voltage sags.

Voltage source inverter with constant dc-link need large energy storage elements. As a result, they aren't suited for applications that require a lot of weight or volume.

Table 1 Comparison of DVR topologies

DVR topologies	Merits	De-merits
Topologies—source side connected converter (no storage)	• Saving since no energy storage • Longer sags can be compensated • More suitable for strong electrical grids	• Uncontrollable dc-link voltage • Limited compensation in deep sags • Draw much line current during the fault
Topologies—load side connected converter (no storage)	• Economical, modular design and compact since no internal energy storage • DC link voltage constant	• Handle larger currents by the series converter • Shunt converter draw nonlinear currents which disrupt the load • Residual supply provide boost energy
Topologies with energy storage—variable DC link	• Converter ratings are less • Less strain on the grid connection • Simple converter topology	• Expensive • DC link voltage is not constant • During severe sags power converter rapidly enters into over modulation
Topologies with energy storage—constant DC link	• Performance improvement because of stored internal energy • Strain is less on grid • Unchanged current from the supply • Complexity of control is less	• Expensive due to energy storage element and converter rating • DC link maintained to be constant • The ability to compensate declines as the stored energy drains away
Transformer less DVR	• Cost, size, and weight reduction • There are no issues with transformer saturation or inrush current	• High voltage applications are not recommended
DVR with AC/AC converter	• No DC link capacitor • No storage element • Size less and compensate long duration sag	• Complexity in control • Bi-directional switches • Deep sag not suited
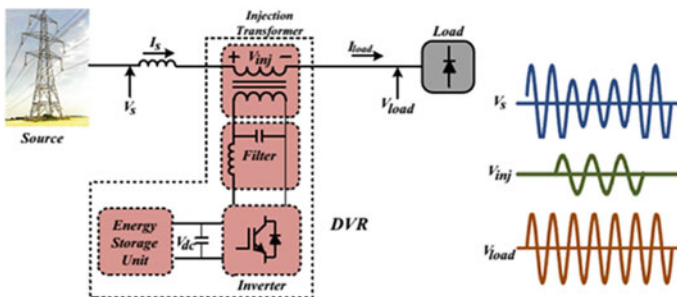
Fig. 1 DVR topology with energy storage

Bi-directional topologies minimize system size and increase efficiency and performance by connecting system and energy storage devices. Bi-directional converters are divided into two main categories, namely, non-isolated and isolated configurations. Non-isolated BDCs are less complicated and more efficient. Many applications, however, need galvanic isolation [3]. The selection of energy storage and BDC in DVRs are analyzed further.

3.1   Energy Storage Element

Energy storage systems finds its application in grid stabilization and power quality enhancements. Batteries, flywheels, fuel cell, ultracapacitor, and superconducting energy storage systems are all viable storage options as discussed in Table 2.

3.2   Power Converter with Energy Storage

Unidirectional dc-dc converter based DVRs achieve power flow in one direction but bidirectional DC-DC converter facilitate energy exchange in both directions. Hence sag power is injected to grid and swell condition power flows from grid to storage element [5]. Power converter used in DVRs are discussed in Table 3.

4   Comparison of DVR with Energy Storage and Converter

From the survey, it is observed that energy storage elements such as flywheel, battery storage, SEMS, super capacitor and ultracapacitors with suitable converter are used in DVR to mitigate the power quality issues. It has been summarized in Table 4.

In this survey, buck-boost BDC converter used with UCAP and bridge type chopper is used for SEMS, fuel cell with multilevel inverter, flywheel with matrix converter, battery [12] with isolated and non-isolated BDC and impedance source converters with PV system to have a better performance in DVRs. The efficiency of the DVR can be still increased by improving the selection of energy storage and reduced rating of power converter [13].

5   Conclusion

The rapid growth of renewable energy, combined with the meteoric rise of non-linear loads, is posing major problems to the quality of traditional unidirectional power flow, which must be addressed. The custom power devices are the best solution and

Table 2 Short-term energy storage technologies
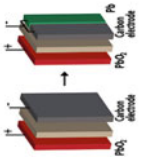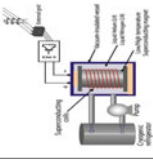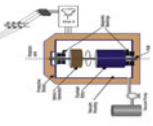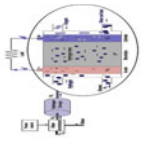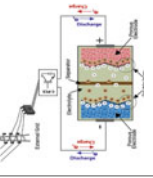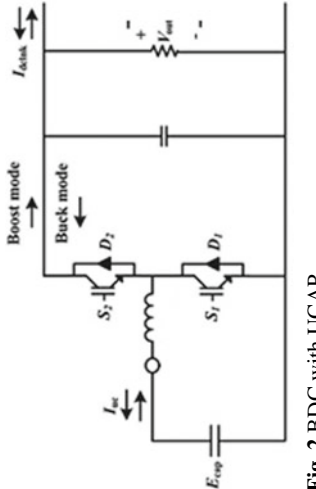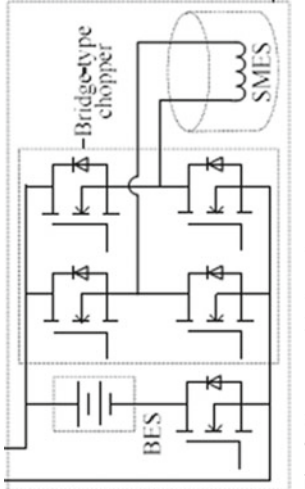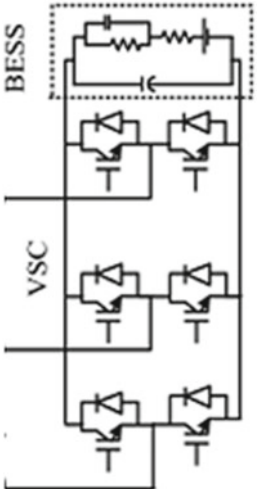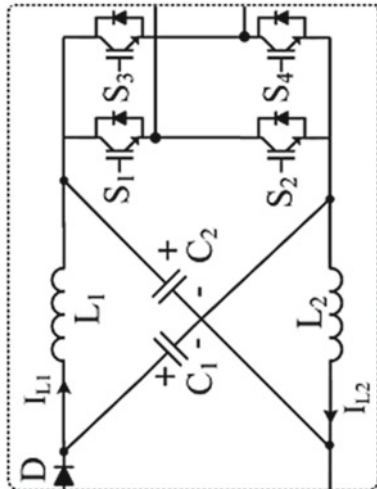
ESS	Battery	SMES	Flywheel	Fuel cell	Ultracapacitors
Module					
Merits	Cost-effective High energy capability	Fast response and efficient (95%)	Modular designs are possible Cost effective	Discontinuous performance Cost of raw material low	High-power density Higher number of charge/discharge cycles Higher terminal voltage
Demerits	Limited life cycle Long duration operation at high power not possible Toxic gas generation Disposal problem	High material and manufacturing cost Cooling arrangement Low-energy density Needs operating temperature below −195.79 °C	Rotational energy losses High frictional losses	Less efficiency Hydrogen storage High cost	Low-energy density

Table 3 Power converters in DVR

Converter	Diagram	Remarks
Bi-directional buck-boost converter with UCAP [4]	Fig. 2 BDC with UCAP	• The inverter and bi-directional dc–dc converter efficiency is around 95% and 92% for sag and 83.3% and 86% for swell, respectively
Bridge type chopper with SMES and battery [6]	Fig. 3 Chopper with SMES	• Efficiency values of the bridge-type and conventional choppers are about 0.876 and 0.526 respectively

Table 3 (continued)

Converter	Diagram	Remarks
Voltage source converter with BESS [7]	 **Fig. 4** VSC with BESS	• Cost of an energy source at the dc bus is high
Impedance source converter [8]	 **Fig. 5** Z source converter	• No observation of inrush current during dynamic change • Discontinuous input current

Table 3 (continued)

Converter	Diagram	Remarks
Dual active bridge converter [5]	**Fig. 6** DAB with battery	• The proposed DVR can achieve the compensation of 30% voltage sag and 25% voltage swell of nominal voltage

Table 4 Comparison of DVR with storage element and suitable converter

Paper	Energy storage element	Level and rating	Sag detection technique and control	Compensation strategy	Converter	Remarks
Jayaprakash et al., 2014 [7]	BESS DC voltage − 300 V	kW Load—10 kVA, 0.8 pf Supply—415 V, 50 Hz	Synchronous reference frame theory (SRF)	In-phase Pre sag Energy optimization	Voltage source converter (VSC)	• DVR's minimum rating is achieved in in-phase technique, but energy source cost is high • Compared to in-phase injection, the voltage injected in Energy optimization is higher
Zixuan Zheng et al., 2018 [9]	SMES + battery (lead acid) 15 * mH, 100A + 48 V (4 * 12)/75 Ah	kW Load—1.4 kW	RMS & d-q detection	Pre-sag	Bi-directional voltage source converter (VSC) and chopper	• Expensive capital costs and equipped refrigeration devices • Improving millisecond-level transient voltage quality • Pre-sag technique causes less transients • d-q control suppress peak voltage and thus protect

Table 4 (continued)

Paper	Energy storage element	Level and rating	Sag detection technique and control	Compensation strategy	Converter	Remarks
Gambôa, Elsevier, 2019 [10]	Flywheel Stored energy—144 kJ	kW Load—110 V (rms), 2 kW	Predictive control	In-phase	AC–AC matrix converter	• All critical loads compensations are met • Solve problem due to coupling in matrix converters
Darvish Falehi, Soft Computing Springer, 2018 [11]	Fuel cell	kW	Adaptive control	Pre-fault compensation strategy	TSBC (boost converter) + QMLI (multilevel inverter)	• Step-up QMLI generate stair-case sine wave not taking the account of number of switches • DC/DC converter reduce switch stress and current ripple

(continued)

Table 4 (continued)

Paper	Energy storage element	Level and rating	Sag detection technique and control	Compensation strategy	Converter	Remarks
Zheng et al., 2017 [6]	HES (SMES + BESS-VRLA) 3.25 mH/240 A + 40 V/100 Ah	kW Load—5 kW Source voltage 380 V, 50 Hz	RMS	Pre-sag	bridge-type chopper and a series-connected MOSFET	• Combines the benefits of quick response, high-power density from SMES and low capital cost, high energy density from BES • In the exclusive BES-based scheme, an initial discharge time delay and a rushed discharging current are avoided • By incorporating both system, the BES system's peak current and power requirements were reduced
Somayajula and Crow 2015 [4]	Ultracapacitor 3 units of 48 V, 165F UCAP DC link voltage – 260 V	kW Supply—208 V, 60 Hz	dq control	In phase	Bidirectional DC/DC converter + VSC	• Independently compensate temporary voltage sags compensate sag last from 3 s to 1 min

among that DVR has gained significant popularity. In this paper, DVR topologies, different energy storage element used in DVR are compared and power converters used to charge and discharge the energy storage element are reviewed. Energy storage topology can compensate deep voltage sags independent of Grid and storage elements are compared for its power and energy density.

References

1. Woodley NH, Morgan L, Sundaram A (1999) Experience with an inverter-based dynamic voltage restorer. IEEE Trans Power Delivery 14:1181–1186
2. Nielsen JG, Blaabjerg F (2005) A detailed comparison of system topologies for dynamic voltage restorers. IEEE Trans Ind Appl 41, 1272–1280
3. Gorji SA, Sahebi HG, Ektesabi M, Rad AB (2019) Topologies and control schemes of bidirectional DC–DC power converters: an overview. IEEE Access 7:117997–118019
4. Somayajula D, Crow ML (2015) An integrated dynamic voltage restorer-ultracapacitor design for improving power quality of the distribution grid. IEEE Trans Sustain Energy 6(2)
5. Inci M et al (2014) The performance improvement of dynamic voltage restorer based on bidirectional dc–dc converter. Springer
6. Zheng ZX et al (2017) Design and evaluation of a mini-size SMES magnet for hybrid energy storage application in a kW-class dynamic voltage restorer. IEEE Trans Appl Super Cond 27(7)
7. Jayaprakash P, Singh B, Kothari DP, Chandra A, Al-Haddad K (2014) Control of reduced-rating dynamic voltage restorer with a battery energy storage system. IEEE Trans Ind Appl 50(2)
8. Sajadian S, Ahmadi R (2018) ZSI for PV systems with LVRT capability IET. IET Renew Power Gener 12(11):1286–1294
9. Gee AM, Robinson F, Yuan W (2017) A superconducting magnetic energy storage-emulator/battery supported dynamic voltage restorer. IEEE Trans Energy Convers 32(1)
10. Gambôa P, Silva JF, Pinto SF, Margato E (2018) Input–output linearization and PI controllers for AC–AC matrix converter based dynamic voltage restorers with flywheel energy storage: a comparison. Electr Power Syst Res 169:214–228
11. Darvish Faleh A et al (2019) Optimal control of novel fuel cell-based DVR using ANFISC-MOSSA to increase FRT capability of DFIG-wind turbine. Soft Comput 23(15)
12. Yan L, Chen X, Zhou X, Sun H, Jiang L (2018) Perturbation compensation-based non-linear adaptive control of ESS-DVR for the LVRT capability improvement of wind farms. IET Renew Power Gener 12(13)
13. Wang J et al (2019) A novel Dual-DC-Port dynamic voltage restorer with reduced-rating integrated DC-DC converter for wide-range voltage sag compensation. IEEE Trans Power Electron 34(8)

A Higher-Order Sliding Mode Observer for SOC Estimation with Higher-Order Sliding Mode Control in Hybrid Electric Vehicle

Prasanth K. Prasad and P. Ramesh Kumar

Abstract The purpose of this study is to drive a fully active hybrid energy storage system (HESS) based on state of charge (SOC). A sliding mode control is used to estimate the SOC of the source such as battery and supercapacitor (SC) used in hybrid electric vehicles. For effective performance of HESS, higher-order sliding mode control method is implemented then the SOCs of battery and supercapacitor are estimated using a super-twisting technique. The advantages of the higher-order sliding mode control are that it can efficiently reject matched system disturbances and remove chattering induced in the control input in short span of time. In order to generate the reference current for effective tracking, rule-based strategy is used. The simulation results shows the estimating method's efficacy by providing accurate estimation of source parameters.

Keywords Electric vehicles (EV) · Hybrid energy storage system (HESS) · State of charge (SOC) · Higher-order sliding mode observer

1 Introduction

Energy consumption has increased considerably in the modern time as a result of the emergence in lifestyle and transportation. In the transportation sector, many of the countries rely on fossil fuel as their principal source of energy [1].

Hybrid electric vehicles (HEVs) are apparent remedy to the global energy dilemma since they can provide green infrastructure while reducing resource usage and reducing carbon emissions [2, 3]. The main snag of development is a strong energy storage system which governs the performance and range of the hybrid electric vehicles. During the driving phase, energy is released to the load, and during the charging period, energy is consumed by the sources.

P. K. Prasad (✉) · P. Ramesh Kumar
Department of Electrical Engineering, Government Engineering College Thrissur, Affiliated to APJ Abdul Kalam Technological University, Thrissur, Kerala, India
e-mail: prasanthkprasad@gectcr.ac.in

Fig. 1 Schematic diagram
for fully active HESS
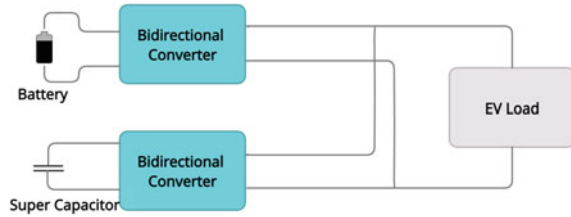
Since the energy storage unit in HEVs is subjected to frequent charging and discharging processes, unintentional battery degradation happens, leading to a diminution in battery life span [4, 5]. In contrast to the battery, storage devices with high-energy like the super capacitor (SC) have much longer life by coupling SC with the battery. Because of storage technical restrictions, it is frequently essential to improve the transient performance and steady-state performance of a storage system; hence, a hybrid energy storage systems (HESSs) are developed. The HESS is made up of two components: (1) high-power storage (HPS) (2) high-energy storage (HES). HPS will absorb/deliver power during transient and peak conditions and thus long-term energy needs can meet by the HESS [6–9]. To assure the battery's safety, HESS utilizes the SC's capability, that includes a comparatively high density of power with fast response time [10]. Essentially, a HESS is designed based on five interconnected factors: storage technology, rated capacity, topology of power converter, energy management strategy, and control methodology, all of them must be properly reviewed. By proper design, HESS is recognized as most efficient energy sources for HEVs [11].

It is possible to classify HESS into three categories: (1) passive HESS, (2) semi-active HESS (3) fully active HESS. The passive type of HESS interconnects sources to the DC bus and aligns the battery and SC in parallel [12]. Passive HESS has a simple architecture and is easy to apply in electric vehicles. By reducing the number of DC-DC converters for an EV system, substantial cost savings can be achieved. In terms of economical aspect, passive type HESS is a viable solution.

In semi-active, HESS uses single DC-DC converter to manage one of its two power sources, while the other power source is linked directly to the DC bus. As a result, the HESS is decoupled to some extent. When compared to passive HESS, this architecture provides superior performance and an excellent cost-performance ratio [13, 14].

In fully active HESS, bidirectional DC/DC converters actively manage the flow of power between both the battery and the SC [15, 16]. This topology structure provides a great degree of controllability since the sources are totally dissociate from the DC bus. This architecture can contribute to the effectiveness, battery life, and DC bus stability by using a properly developed control technique. Figure 1 shows a fully active HESS.

HESS relies on power distribution and energy management technologies. An energy management system (EMS) platform is often used to build energy management strategies. The EMS performs the following tasks:

- Regular monitoring of power flow.
- Providing reference signals for controllers to manage power flow.
- Manage the system's relative state of charge in order to achieve overall vehicle set objectives like fuel economy and performance.

As a result, the overall system's power output capacity is improved, the battery's life is increased, and system replacement costs are reduced. In order to facilitate effective power sharing amongst hybrid sources a powerful energy management system (EMS) should be established.

EMS has been employed in a variety of ways to protect the associated sources. Commonly using energy management techniques can be listed as, firstly, a global optimization approaches based on Poynting's minimum principle [9], dynamic programming (DP) [17], and real-time approaches like filtration-based strategy [18], rule-based technique [19], model predictive method (MPC) [20], "all or nothing" strategy [21], fuzzy logic approaches [5]. Above-mentioned management techniques use optimal power demand allocation to generate battery and SC current references. Among the EMS listed, rule-based methodology is a simple and efficient online energy management technique based on system parameters like power demand, SOCs of sources.

For optimal EMS operation, SOC of battery and SC must be known. The significance of SOC is that it informs how long the battery can operate before it has to be charged or replaced, i.e. it expresses the HEV's driving range. Also, With more accurate SOC estimation, the energy management controller can make greater use of the of the HESS without risking harmful overcharge/overdischarge or other catastrophic threats. Since the SOC of a source cannot be measured physically from a system, alternative method for determining the SOC using system states must be used.

The SOC may be estimated using a multitude of approaches. One of the classical methods of SOC estimation is open-circuit voltage (OCV) method which rely on the linear relation between OCV and SOC. The merits of the OCV technique are its simplicity of use and better degree of accuracy. The flaws are also clear; to measure the terminal voltage, it must be positioned for an extended amount of time, making it impossible to use in on-line estimating [22]. The most popular method for estimation is Coulomb counting [23], which integrates the current across time. Merits are easy computation, a reliable method, and real-time measurement will be provided. Because the quantity to be measured is so vast that it cannot be measured in a timely manner, there is always the possibility of inaccuracy, also, ageing factor is not considered in this method. Finally, most significant disadvantage is the sensitivity to sensor accuracy problems, which tend to accumulate as the current is integrated.

In modern methods based on control theory, some sophisticated methods have arisen in past few decades, such as Kalman filtering [24], neural network method, and sliding mode observer [25], which offer online real-time estimate and have emerged as a distinct research hotspot. The Kalman filter method relies on optimum design of the least-variance of the state of the power system. The benefit is that it eliminates the Coulomb counting method cumulative inaccuracy over time as well as it does not place a significant value on the correctness of the initial SOC. The limitation of

accuracy is heavily reliant on the formulation of a equivalent model of battery and error can induce due to non linearity of the model, due to threat of disturbances and the time dependency of model. In recent research, good amount of modification is done in Kalman filter method like Extended Kalman Filter method (EKF), which linearise the nonlinear system, Unscented Kalman Filtering approach (UKF), which handles probability distribution along with non linear problem, Central Difference Kalman Filtering (CDKF) method, which is complaining central difference method with KF. Another type of modern approach is neural network method which is a stream of artificial intelligent. The benefits of this technique are that, it can rapidly estimate the SOC, as well as the simultaneous and universal searching techniques, fast, it has precision and greater convergence speed in the trial. Downsides can be listed as requires a significant number of training data as reinforcement to finish the training system, also too complicated and takes a huge amount of computation.

Sliding mode observer (SMO) is the another prominent topic of research based on control theory. One of the distinctive features is the sliding motion on the error between the observer's output and measured plant output and delivers the estimates of state that are precisely proportional to the plant's actual output. The estimated state is mapped into the hyperplane, yielding an estimated output that is identical to the measured output, and the process on error convergence will ensures in finite time. This approach provides a high level of robustness against model uncertainty and external disturbances [26].

For the controller, with a conventional PID controller, high reliability can be ensured and relatively easy to implement. In the case of local linearization, the performance may degrade as a result of varying functioning states [27]. Hence, a nonlinear method of control approaches is proposed to resolve the issue [28–30]. A sliding mode control based on conventional approach along with a Lyapunov function is suggested in [28] to track the reference which is generated by EMS and to regulate the bus voltage. In [29], adaptive-SM control is put forward which rely on estimators, that can predict discrepancy in load and disturbances by Lyapunov-based function and properly designed observers. The control problem is designed into a problem of numerical method optimization with linear-matrix inequality constraints in [30]. Feedback law based on two-state is implemented in linear-matrix inequality to regulate voltage and current.

In sliding mode control (SMC), when the system's states are confined in the sliding surface, in a finite time, the state convergence will occur and rejects the matched bounded disturbances. The limitations are, the occurrence of discontinuous control action, which can leads to a chattering which is a phenomena characterized by oscillations at the system's output that can be hazardous to control systems. To alleviate the phenomenon of chattering, a higher-order SMC (HOSMC) can be used [31]. HOSM may be realized using a variety of algorithms. In specifically, second-order sliding mode algorithms can be find by terminal SMC, sub-optimal controller, the twisting algorithm-based controller, and the super-twisting (ST) algorithm-based controller. In peculiar, the twisting algorithm enforces the sliding variable S of relative degree two, but necessitating knowledge of Ṡ. Although, Ṡ is not required by the super-twisting algorithm, and also the sliding variable has relative degree one. Hence,

the ST algorithm is now preferred over the conventional SM controllers since it eliminates chattering.

In this paper, a uniform continuous HOSMC (UCHOSMC) is used. This hybrid controller combines a super-twisting controller to achieve disturbance rejection with a controller meant for uniform convergence under no disturbances [32]. And super-twisting observer is used for the accurate estimation of the state.

Here is how the remaining portion of the paper is organized: Sect. 2 defines proposed system model. The HESS system and model equation are addressed in Sect. 3. Section 4 investigates RC model for SOC estimation. In Sect. 5, rule-based energy management system for HESS is explained. Section 6 shows SOC estimation. Section 7 gives a description about the controller that used for the work. Section 8 gives simulation results for the system. Finally, Sect. 9 concludes the work by giving the results and inference from the work.

2 Proposed System Model

Figure 2 explains system model. The system is comprised of the HESS system, the EV system, and the estimation and control mechanism. The HESS system consists of two sources, namely the battery and the SC, a bidirectional converter coupled with the sources to provide charging and discharging purpose of the sources. The inverter, driving motor, and transmission system are included in an EV system. The third section incorporates the controlling mechanisms, which comprise an estimator to forecast source SOCs, after that, an EMS to generate reference currents for the controller, then a controller to track the reference currents, finally a switching pulse generator to generate switching pulses for the converter switches. The light coloured lines represent the power flow in the system, while the dark coloured lines indicate the system's control signals.
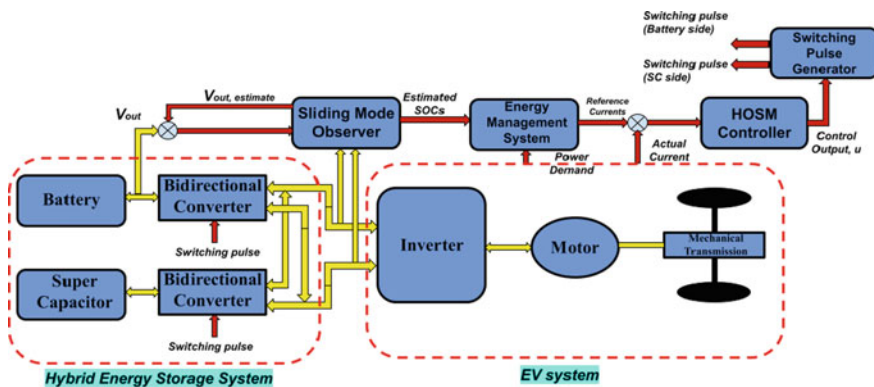
Fig. 2 Proposed model of hybrid electric vehicle

Fig. 3 Fully active HESS
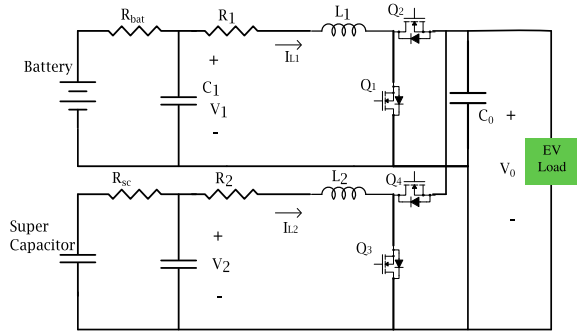
3 HESS Modelling

Figure 3 illustrates the fully active HESS architecture. To deliver average power to the load, battery is used as the power source and is coupled to the bus via a bidirectional DC-DC converter. The SC is connected through a bidirectional DC-DC converter to the bus which is meant to provide peak power. An inverter connects the HESS output to the load. The switches S_1 (S_3) and S_2 (S_4) will not be operate at the same time. D_1, D_3 are the duty ratios during on state of the switches S_1 and S_3, respectively.

Table 1 illustrates the terminology used in Fig. 3.

The average model for HESS expressed as [14],

$$\dot{V}_1 = -\frac{V_{C_1}}{R_{bat}C_1} - \frac{i_{L_1}}{C_1} + \frac{V_{bat}}{R_{bat}C_1} \tag{1}$$

$$\dot{V}_2 = -\frac{V_{C_2}}{R_{sc}C_2} - \frac{i_{L_2}}{C_2} + \frac{V_{sc}}{R_{sc}C_2} \tag{2}$$

$$\dot{i}_{L_1} = \frac{V_{C_1}}{L_1} - i_1\frac{R_{L1}}{L_1} - \frac{V_0}{L_1} + V_0\frac{D_1}{L_1} \tag{3}$$

$$\dot{i}_{L_2} = \frac{V_{C_2}}{L_2} - i_2\frac{R_{L2}}{L_2} - \frac{V_0}{L_2} + V_0\frac{D_3}{L_2} \tag{4}$$

$$\dot{V}_0 = \frac{i_{L_1} + i_{L_2}}{C_0} - \frac{i_m}{C_0} - D_1\frac{i_{L_1}}{C_0} - D_3\frac{i_{L_2}}{C_0} \tag{5}$$

State space model of HESS according to the averaged model can be expressed as,

$$\dot{x} = (A_0 + A_1D_1 + A_2D_3)x + B_ev_e, \qquad y = Cx \tag{6}$$

In (6), x indicates the state variables of the system and y indicates the output, and, $A_0, A_1, A_2, B_e, C, v_e$ are the parameter matrices [28]. The HESS model is substantiated in [29].

Table 1 HESS terminologies

Notation	R_{bat}	R_{sc}	C_1
HESS parameters	Battery parasite resistance	SC parasite resistance	Filter-capacitor across battery
Notation	L_2	i_{L1}	i_{L2}
HESS parameters	Series inductance of SC	Current through L_1	Current through L_2
Notation	V_1	V_2	L_2
HESS parameters	Voltage across C_1	Voltage across C_2	Series inductance of battery
Notation	R_2	V_{bat}, V_{sc}	C_0
HESS parameters	Series resistance of L_2	Battery and SC voltage	DC bus capacitance
Notation	C_2	R_1	V_0
HESS parameters	Filter-capacitor across	Series resistance of L_1	Bus voltage

$$A_0 = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & W_0 \end{bmatrix} \quad A_1 = \begin{bmatrix} 0 & 0 \\ 0 & W_1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & W_2 \end{bmatrix}$$

where

$$A_{11} = \begin{bmatrix} -\frac{1}{R_{\text{bat}}C_1} & 0 \\ 0 & -\frac{1}{R_{\text{sc}}C_2} \end{bmatrix} \quad A_{12} = \begin{bmatrix} -\frac{1}{C_1} & 0 & 0 \\ 0 & -\frac{1}{C_2} & 0 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} \frac{1}{L_1} & 0 \\ 0 & \frac{1}{L_2} \\ 0 & 0 \end{bmatrix}, \quad W_0 = \begin{bmatrix} -\frac{R_{L1}+R_{\text{on}2}}{L_1} & 0 & -\frac{1}{L_1} \\ 0 & -\frac{R_{L2}+R_{\text{on}4}}{L_2} & -\frac{1}{L_2} \\ \frac{1}{C_0} & \frac{1}{C_0} & -\frac{i_m}{V_0 C_0} \end{bmatrix}, \quad W_1 = \begin{bmatrix} \frac{R_{\text{on}2}-R_{\text{on}1}}{L_1} & 0 & \frac{1}{L_1} \\ 0 & 0 & 0 \\ -\frac{1}{C_0} & 0 & 0 \end{bmatrix},$$

$$W_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{R_{\text{on}4}-R_{\text{on}3}}{L_2} & \frac{1}{L_2} \\ 0 & -\frac{1}{C_0} & 0 \end{bmatrix}, \quad v_e = \begin{bmatrix} V_{\text{bat}} \\ V_{\text{sc}} \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad B_e = \begin{bmatrix} \frac{1}{R_{\text{bat}}C_1} & 0 \\ 0 & \frac{1}{R_{\text{sc}}C_2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

4 RC Model for SOC Estimation

4.1 Second-Order RC Equivalent Model of Battery

The prime aim is to design an accurate battery model, to improve the accuracy of SOC estimation. Electrochemical and equivalent circuit models are two types of bat-

Fig. 4 Battery RC model
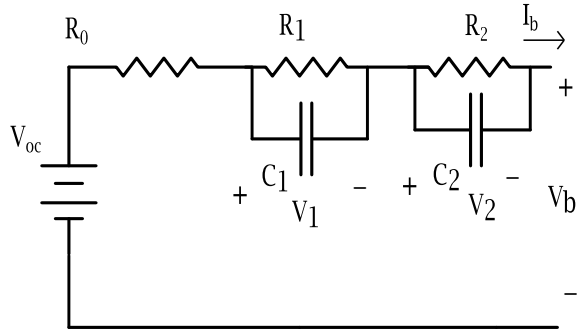
Table 2 Battery model RC model parameters

Terms	Term description
R_0	Ohmic resistance
R_1	Resistance due to activation polarization
C_1	Capacitance due to activation polarization
C_2	Concentration polarization capacitance
V_1 and V_2	Terminal voltages of C_1 and C_2
V_b	Terminal voltages
I_b	Current of the battery

tery models that are extensively used to simulate the external features of a battery. Electrochemical models are made up of a collection of coupled partial differential equations that characterize how the potential of the battery is developed and altered by electrochemical processes within the battery. It is somewhat precise, but its computation is very complicated [33]. The equivalent circuit model, which is built with fundamental circuit components including resistors, capacitors, and voltage sources, maintains a compromise between intricacy and correctness [34].

A second-order RC equivalent model is assessed, as in Fig. 4. The model parameters are listed in Table 2.

V_b indicates terminal voltage and I indicates the current. By applying Krichhoff's law V_b can given as,

$$V_b = V_{oc} - V_1 - V_2 - R_0 I \tag{7}$$

The differential equations for SOC of battery, V_1 and V_2 can be represented as,

$$\dot{S}_{SOC} = -\frac{I}{Q} \tag{8}$$

where S_{SOC} is the SOC of battery.

$$\dot{V}_1 = -\frac{V_1}{R_1 C_1} + \frac{I}{C_1} \quad \dot{V}_2 = -\frac{V_2}{R_2 C_2} + \frac{I}{C_2} \tag{9}$$

By using piecewise linearization method, open-circuit voltage (OCV) is represented as linear function of SOC, and given as,

$$V_{OC} = k_0 S_{SOC} + k_1 \tag{10}$$

where k_0 and k_1 are constants.

By assuming I is constant and substituting (8) to (10) in (7) will yields,

$$\dot{V}_b = -k_0 \left(\frac{I}{Q_n}\right) + \frac{V_1}{R_1 C_1} - \frac{I}{C_1} + \frac{V_2}{R_2 C_2} - \frac{I}{C_m} + \zeta \tag{11}$$

A term ζ added to the equation which represents the error during modelling and uncertain disturbances.

From (7) obtaining I and substituting in (8) and rearranging (9) will gives system's state equation:

$$\dot{V}_b = -a_1 V_b + a_1 k_0 \text{SOC} - a_4 V_1 - a_3 V_2 - b_1 I + a_1 k_1 \tag{12}$$

$$\dot{S}_{SOC} = a_2 V_b - a_2 k_0 \text{SOC} + a_2 V_1 + a_2 V_2 - a_2 k_1 \tag{13}$$

$$\dot{V}_1 = -a_3 V_1 + b_2 I \tag{14}$$

$$\dot{V}_2 = -a_4 V_2 + b_3 I \tag{15}$$

where $a_1, a_2, a_3, b_1, b_2, b_3$ are the constants from system parameters which can obtained as,

$$b_1 = \frac{k_0}{Q_n} + \frac{R_0}{R_1 C_1} + \frac{R_0}{R_2 C_2} + \frac{1}{C_1} \quad b_2 = \frac{1}{C_1} \quad b_3 = \frac{1}{C_2} \tag{16}$$

$$a_1 = \frac{1}{R_1 C_1} + \frac{1}{R_2 C_2} \quad a_2 = \frac{1}{R_0 Q_n} \quad a_3 = \frac{1}{R_1 C_1} \quad a_4 = \frac{1}{R_2 C_2} \tag{17}$$

The state space model can be expressed as,

$$\dot{x} = A_{\text{bat}} x + B_{\text{bat}} u + \Phi + \Delta z, \quad y = C_{\text{bat}} x \tag{18}$$

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$ are the state vector and output vector of the battery model. A_{bat}, B_{bat}, C_{bat}, D_{bat} are the matrices formed by the parameters, Φ and Δz are column matrices, $a_1, a_2, a_3, b_1, b_2, b_3$, which can expressed as,

$$A_{\text{bat}} = \begin{bmatrix} -a_1 & a_1 k_0 & -a_4 & -a_3 \\ a_2 & a_2 k_0 & a_2 & a_2 \\ 0 & 0 & -a_3 & 0 \\ 0 & 0 & 0 & -a_4 \end{bmatrix}, \quad B_{\text{bat}} = \begin{bmatrix} -b_1 \\ 0 \\ b_2 \\ b_3 \end{bmatrix},$$

$$\Phi = \begin{bmatrix} a_1 k_1 \\ -a_2 k_1 \\ 0 \\ 0 \end{bmatrix}, \quad \Delta z = \begin{bmatrix} \zeta \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad C_{\text{bat}}^T = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (19)$$

In this work, the primary goal is to develop an sliding mode observer capable of estimating battery SOC using the model provided.

4.2 SC RC Model

In the literature, the most prevalent equivalent models of SCs are shown in Fig. 5. The equivalent circuit model of the super capacitor is developed to demonstrate the charging and discharging properties of super capacitor.

In Fig. 5, R_{sc1} indicates the equivalent resistance and C_{sc1} denotes the capacitance of large pores. R_{sc2} represents charge redistribution resistance, while C_{sc2} indicates capacitance due to charge redistribution, of small pores. Also, R_L gives self discharge of SC [35].

In SC, charge will stored in small pores as well as in large pores. Hence, stored charge in a capacitor can given as,

$$Q_{\text{sc}} = C_{\text{sc1}} V_{\text{sc1}} + C_{\text{sc2}} V_{\text{sc2}} \qquad (20)$$

where V_{sc1} and V_{sc2} indicates the voltage across capacitors C_{sc1} and C_{sc2}, respectively. In super capacitors, rated capacity can be derived as,

$$Q_{\text{sc,rated}} = (C_{\text{sc1}} + C_{\text{sc2}}) V_{\text{rated}} \qquad (21)$$

Fig. 5 SC RC model
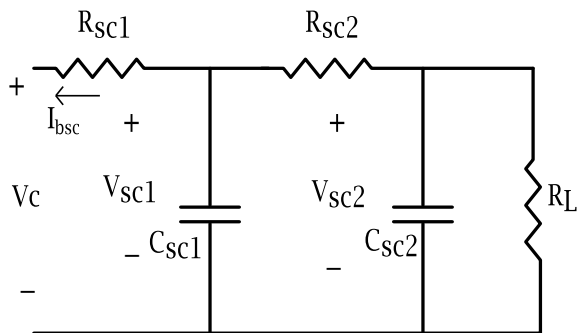
where, V_{rated} is the rated voltage of the SC. Therefore, SOC of SC can expressed as,

$$\text{SOC} = \frac{Q_{\text{sc}}}{Q_{\text{sc,rated}}} = \frac{C_{\text{sc1}} V_{\text{sc1}} + C_{\text{sc2}} V_{\text{sc2}}}{(C_{\text{sc1}} + C_{\text{sc2}}) V_{\text{rated}}} \times 100\% \tag{22}$$

From Fig. 5 of SC RC model state model equation of SC can be given as:

$$\dot{V}_{\text{sc1}} = -\frac{1}{R_2 C_1} V_{\text{sc1}} + \frac{1}{R_2 C_1} V_{\text{sc2}} + \frac{1}{C} I_b, \qquad \dot{V}_{\text{sc2}} = -\frac{1}{R_2 C_2} V_{\text{sc1}} + \frac{1}{R_2 C_2} V_{\text{sc2}} \tag{23}$$

Also, the terminal voltage V_t can written as,

$$V_c = I_b R_{\text{sc1}} + V_{\text{sc1}} \tag{24}$$

From the estimation of states V_{sc1} and V_{sc2}, SOC can given determined using (22).

5 Energy Management System

The EMS technique splits the power requirement between the SC and battery based on the load power. A motor will be the load for an electric vehicle, which is powered by an inverter powered by HESS. The battery reference current will generate by EMS algorithm $i_{\text{bat-ref}}$ according to power demand and the SOCs of sources. The condition considered in EMS is that the SOC of the super capacitor is always higher than 50% of maximum SOC.

Here, a rule-based intelligent power management approach has been developed. Without employing additional DC bus and SC voltage controls, the new method derives the reference current quantity of the involved sources. The power management algorithm will determine the system configuration for each working mode.

The power demand from the system will be provided to the EMS, and the battery SOC and SC SOC are approximated using some observer. Initially, EMS will evaluate if the power demand is higher than zero. If demand is greater than zero, EMS will verify whether the power level is greater than p_{min}, which is minimum power defined for the battery to supply during low SOC level. Then the SOC level will be checked; if the demand is more than p_{min}, the SC, SOC will be checked; if the SOC is greater than 50%, p_{min} will be provided by the battery and the remainder by the SC. If the SOC is less than 50%, the battery will deliver the necessary power.

In contrast, if the power demand is less than p_{min}, the battery's SOC level will be checked. If the battery's SOC is more than 20%, the battery will deliver the needed power; otherwise, if the SOC is less than 20%, the battery will offer a specific low power, described as p_{ch}.

In another scenario, the SC will be charged from the load if the power demand is lower than zero, that is in regenerative mode.

6   SOC Estimation

A reliable, efficient, and robust HESS system needs an accurate estimation of SOC in the presence of model uncertainty and noise. Sliding mode observer approach is used here to estimate SOC of battery and SC because of its superior performance in nonlinear scenarios. Classification of SMO can be described as below.

6.1   Super-Twisting Observer

Super-twisting algorithm is one of the approaches in higher-order observer and the relative degree is one. This algorithm can offer the observer with effective elimination of chattering and finite time convergence of the states. This framework incorporates a continuous sliding variable function along with a discontinuous sliding variable.

6.2   Design of the Super-Twisting Sliding Mode Observer for Battery SOC Estimation

The ST observer equation to estimate the SOC can be given as,

$$\dot{\hat{x}}_1 = \hat{x}_2 + z_1 \quad \dot{\hat{x}}_2 = f(\hat{x}, t) + g(\hat{x}, t)u + z_2 \tag{25}$$

(25) can represent in terms of system variables as, $\hat{x}_1$. From (10) $\hat{V}_{oc}$ can be written as,

$$\dot{\hat{V}}_{oc} = k_0 \dot{\hat{S}}_{SOC} \tag{26}$$

From (11) and (23), (24) can able to rewritten as,

$$\dot{\hat{x}}_1 = \hat{x}_2 = k_0 \dot{\hat{S}}_{SOC} \tag{27}$$

The sliding manifold can developed as,

$$e_1 = x_1 - \hat{x}_1.$$

Then the correction term for (25) can designed as

$$z_1 = k_1 \sqrt{|e_1|} \operatorname{sign}(e_1) \quad z_2 = k_2 \operatorname{sign}(e_1) \tag{28}$$

So, by proper selection of $k_1 > 0$ and $k_2 > 0$, the error e_1 converges to zero in time that is $t > T_0$.

Thus from proper selection of constant k_0, SOC of battery can be estimated from (10). Since $\hat{x}_1$ and $\hat{x}_2$ are estimated, value of $\hat{V}_b$ can be find out.

From the error equation $e_1 = x_1 - \hat{x}_1$, by substituting the terms of x_1 and $\hat{x}_1$ will gives,

$$e = (V_{oc} - V_b) - (\hat{V}_{oc} - \hat{V}_b) \tag{29}$$

If error of states defined as,

$$e_1 = S_{SOC} - \hat{S}_{SOC}, \quad e_2 = V_b - \hat{V}_b, \quad e_3 = V_1 - \hat{V}_1, \quad e_4 = V_2 - \hat{V}_2 \tag{30}$$

From (30), the error dynamics can be represented as,

$$\dot{e} = k_0 \dot{e}_1 + a_1 \dot{e}_2 - a_2 k_0 \dot{e}_1 + a_4 \dot{e}_3 + a_3 \dot{e}_4 \tag{31}$$

where, constants k_0, a_1, a_2, a_3 are the constants in matrices A_{bat}, Φ which should me positive matrices. By proper selection of constants the matrix conditions will be satisfied, and error in (31) will be converges to zero in finite time period.

6.3 Design of the Super-Twisting Sliding Mode Observer for SC SOC Estimation

To estimate the SOC of SC accurately using ST observer, observer dynamics can be written as in (34) and (35),

$$\hat{x}_{1\_sc} = \hat{V}_{1sc} - \hat{V}_c + z_{1\_sc} \quad \dot{\hat{x}}_{1\_sc} = \dot{\hat{V}}_{1sc} - \dot{\hat{V}}_c + z_{2\_sc} = \hat{x}_{2\_sc} \tag{32}$$

By using (24), (32) can modified as,

$$\dot{\hat{x}}_{1\_sc} = -I_b R_{sc1} + z_{2\_sc} = \hat{x}_{2\_sc} \tag{33}$$

The sliding manifold for the system can be designed as,

$$e_{sc} = V_c - \hat{V}_c$$

The correction term for estimation algorithm can developed as,

$$z_{1\_sc} = k_3 \sqrt{|e_{sc}|} \text{sign}(e_{sc}) \quad z_{2\_sc} = k_4 \text{sign}(e_{sc}) \tag{34}$$

By using (34), the error e_{1sc} will converge to zero and the states mentioned in (32) can be estimated.

By taking the known value of terminal voltage V_t of SC and from (25) V_{1sc} can be find out. Considering the errors as,

$$e_{2sc} = V_{sc1} - \hat{V}_{sc1} \tag{35}$$

Then error dynamics for the system then written as,

$$\dot{e}_{1sc} = k_{0sc}\dot{e}_{1sc} + k_{0sc}\dot{e}_{2sc} \tag{36}$$

where, $k_{0sc} = \frac{1}{R_2 C_1}$ and ensures finite time convergence.

To estimate $\hat{V}_{2sc}$ writing the observer dynamics as,

$$\hat{x}_{3\_sc} = \hat{V}_{2sc} - \hat{V}_c + z_{3\_sc} \quad \dot{\hat{x}}_{3\_sc} = \dot{\hat{V}}_{2sc} - \dot{\hat{V}}_c + z_{4\_sc} = \hat{x}_{4\_sc} \tag{37}$$

The correction term for estimation can taken as,

$$z_{3\_sc} = k_3\sqrt{|e_{sc}|}\text{sign}(e_{sc}) \quad z_{4\_sc} = k_4\text{sign}(e_{sc}) \tag{38}$$

Therefore from the true value of V_t; V_{2sc} can be identified.

After obtaining V_{1sc} and V_{2sc}, by using (22) SOC of SC can be estimated, i.e.,

$$\text{S}\hat{\text{O}}\text{C} = \frac{C_{sc1}\hat{V}_{sc1} + C_{sc2}\hat{V}_{sc2}}{(C_{sc1} + C_{sc2})V_{\text{rated}}} \times 100\% \tag{39}$$

Considering the errors as,

$$e_{3sc} = V_{sc2} - \hat{V}_{sc2} \tag{40}$$

Then error dynamics for the system then written as,

$$\dot{e}_{4sc} = k_{1sc}\dot{e}_{3sc} + k_{1sc}\dot{e}_{3sc} \tag{41}$$

where, $k_{1sc} = \frac{1}{R_2 C_2}$, and error will converge in finite time when properly choosing constants $k_1 > 0$.

7 Controller

A robust and precise method must be implemented in the proposed system to track the reference current obtained from the EMS. Higher-order sliding mode control method will be an excellent option for the tracking of the reference values.

Uniform continuous HOSMC (UCHOSMC) is implemented in this system to deliver finite time state convergences despite initial conditions when contrasted to

Fig. 6 Drive cycle
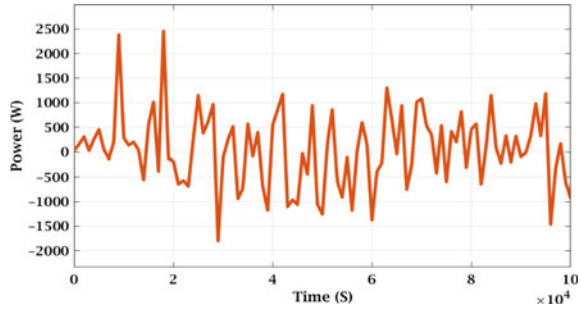
the super-twisting algorithm. UCHOSMC is a combination of couple of controllers: first one is that could uniformly stabilize the system during no disturbances, and other uses a uniform ST algorithm that have a capability of disturbance observation. Under this control scheme, the uniform ST controller takes over the disturbances effect and the uniform controller is responsible for governing the closed loop system's response. Thus, this approach ensures an exact continuous control law and convergence in finite time [32].

8 Simulation Results

The performance is assessed on the MATLAB platform using the observer algorithm. In the simulation, the load is considered to be a current source equivalent to the EV's load.

The bus voltage is regulated to 240 V, the voltages of battery and SC are maintained slightly below 240 V. When the load current hits 8 A, the SC delivers the energy, and the SC gets charged from the load during regenerative braking. The drive cycle used for the simulation of proposed observer in the system is depicted in Fig. 6.

8.1 SOC Estimation

It is assumed in this case both the battery SOC and the SC SOC are 80% charged at the start and are discharging in the allotted period. The simulation result and analysis from the result is described in section below.

8.1.1 SOC Estimation of Battery

The proposed estimation method is modelled and compared with the conventional approaches like, Coulomb counting (CC) method and extended Kalman filter method.

Fig. 7 Battery SOC estimation
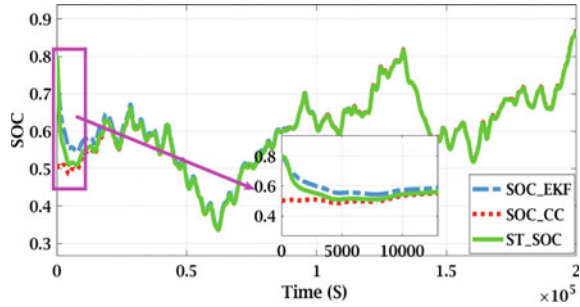
Fig. 8 Error convergence of battery SOC
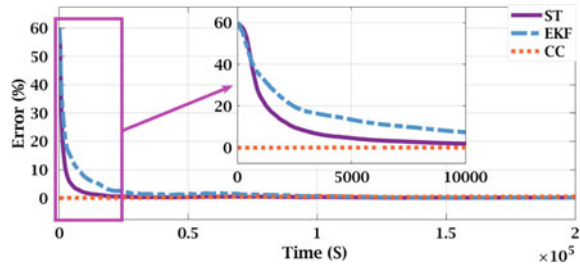
Fig. 9 SOC estimation in various initial conditions of battery
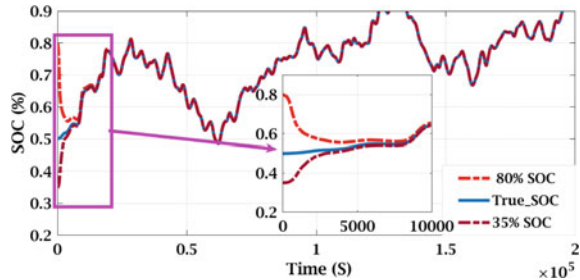
The result obtained is shown in Fig. 7. Since, CC method is nothing but integration approach of current, it is directly following the true SOC value. In the case of ST observer, it is tightly following the true SOC from any value of initial state. When comparing with extended Kalman filter (EKF) estimation, the ST technique reduces the amount of time to reach true value, in ST observer approach, the initial value to its true value in short time as compared with EKF method.

Figure 8 depicts the convergence of error value of SOC on each method. It ensures that the proposed method will provide less time of error convergence to zero value. Also, strongly hold the error value around zero whatever disturbance comes into the system.

Figure 9 indicates the convergence of estimated value of SOC in different initial conditions. Whatever be the initial condition the observer will accurately take the estimation value into the true SOC level in minimum time.

Table 3 RMSE and time taken to converge error into permissible limit for battery

Method	Root mean square error		Error convergence below 10% (s)	
	SOC(0) = 35%	SOC(0) = 80%	SOC(0) = 35%	SOC(0) = 80%
ST observer	0.0598	0.0615	2330	2610
EKF approach	0.0731	0.0692	6940	5705

Fig. 10 SOC estimation of SC
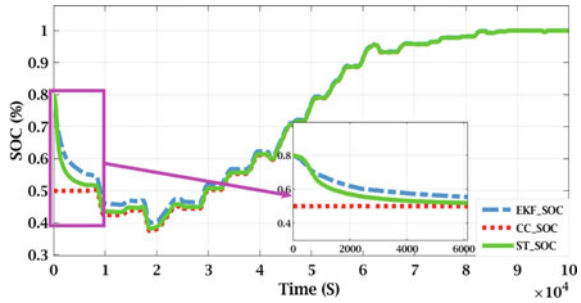
Fig. 11 Error convergence of SOC of SC
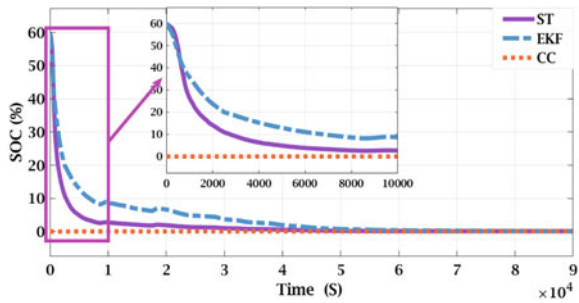
To quantify the enhanced exactness of the STO-based SOC estimation, Table 3 gives the Root Mean Square Error (RMSE) of SOC estimate at different levels also shows the time taken by ST observer and EKF method to take the error below the permissible limit that is taken as 10%.

8.1.2 SOC Estimation of SC

As done in battery SOC estimation simulation and analysis, the SOC of SC is estimated and compared using the ST algorithm, as well as traditional approaches such as Coulomb's counting and extended Kalman filter methods. As according EMS, the SC system will release energy whenever needed and it will charge from the load while in regenerative mode. Figure 10 illustrates the estimated SOC of the SC system by utilizing multiple ways such as the ST algorithm, the extended Kalman filter, and the Coulomb's counting method.

Fig. 12 SOC estimation in various initial conditions of SC
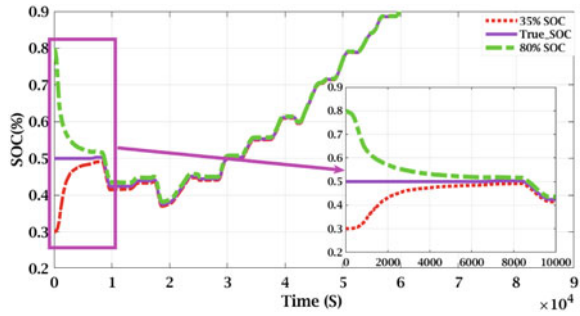
Table 4 RMSE and time taken to converge error into permissible limit for SC

Method	Root mean square error		Error convergence below 10% (seconds)	
	SOC(0) = 35%	SOC(0) = 80%	SOC(0) = 35%	SOC(0) = 80%
ST observer	0.0775	0.0507	2500	2780
EKF approach	0.0879	0.0635	6940	6810

Figure 11 depicts the convergence of error value of SOC on each method. ST observer took less span of time when compared with other approaches and ensures zero error in progress.

Figure 12 shows the observer estimates the SOC to its true value in different initial values in less expenditure of time. After that estimated value strictly following the true value of SOC of SC.

The RSME of estimated SC SOC is given in Table 4, ensuring the robustness of the proposed observer and illustrates the error convergence time for the SC in different SOC conditions to permissible limit.

9 Conclusion

The implementation of a HOSMO for SOC estimation in fully active HESS in EVs is proposed. A energy management system (EMS) based on rule-based strategy effectively regenerates the reference current based on power demand, SOC of battery and SC. To gather source SOCs, a super-twisting observer is used, which can reliably estimate SOCs using its model. When compared to conventional techniques of estimating regardless of initial condition, the suggested algorithm successfully eliminates the disturbance and accurately estimates the state in a small period of time. Also, ensures finite time convergence of error during estimation. Thus, HESS in EVs can be effectively utilized for trustful and efficient operation with diverse load conditions using this strategy.

References

1. Aneke M, Wang M (2016) Energy storage technologies and real life applications—a state of the art review. Appl Energy 179:350–377
2. Martinez CM, Hu X, Cao D, Velenis E, Gao B, Wellers M (2016) Energy management in plug-in hybrid electric vehicles: recent progress and a connected vehicles perspective. IEEE Trans Veh Technol 66(6):4534–4549
3. Bayati M, Abedi M, Gharehpetian GB, Farahmandrad M (2019) Short-term interaction between electric vehicles and microgrid in decentralized vehicle-to-grid control methods. Prot Control Mod Power Syst 4:45–52
4. Vinot E, Trigui R (2013) Optimal energy management of HEVs with hybrid storage system. Energy Convers Manag 76:437–452
5. Song Z, Hofmann H, Li J, Hou J, Han X, Ouyang M (2014) Energy management strategies comparison for electric vehicles with hybrid energy storage system. Appl Energy 134:321–331
6. Hajiaghasi S, Salemnia A, Hamzeh M (2019) Hybrid energy storage system for microgrids applications: a review. J Energy Storage 21:543–570
7. Wang B, Xu J, Wai R, Cao B (2017) Adaptive sliding-mode with hysteresis control strategy for simple multimode hybrid energy storage system in electric vehicles. IEEE Trans Ind Electron 64(64):1404–1414
8. Trovão JP, Pereirinha PG, Jorge HM, Antunes CH (2013) A multi-level energy management system for multi-source electric vehicles—an integrated rule-based meta-heuristic approach. Appl Energy 105:304–318
9. He H, Xiong R, Zhao K, Liu Z (2013) Energy management strategy research on a hybrid power system by hardware-in-loop experiments. Appl Energy 112:1311–1317
10. Wang B, Xu J, Cao B, Zhou X (2015) A novel multimode hybrid energy storage system and its energy management strategy for electric vehicles. J Power Sources 281:432–443
11. Blanes JM, Gutierrez R, Garrigos A, Lizan J, Cuadrado J (2013) Electric vehicle battery life extension using ultracapacitors and an FPGA controlled interleaved buck-boost converter. IEEE Trans Power Electron 28:5940–5948
12. Dougal RA, Liu S, White RE (2002) Power and life extension of battery-ultracapacitor hybrids. IEEE Trans Compon Packag Technol 25(1):120–131
13. Lhomme W, Delarue P, Barrade P, Bouscayrol A, Rufer A (2005) Design and control of a supercapacitor storage system for traction applications. In: Fourteenth IAS annual meeting. Conference record of the 2005 industry applications conference, vol 3, pp 2013–2020
14. Cao J, Emadi A (2012) A new battery/ultracapacitor hybrid energy storage system for electric, hybrid, and plug-in hybrid electric vehicles. IEEE Trans Power Electron 27(1):122–132
15. Camara MB, Gualous H, Gustin F, Berthon A, Dakyo B (2010) DC/DC converter design for supercapacitor and battery power management in hybrid vehicle applications-polynomial control strategy. IEEE Trans Ind Electron 57(2):587–597
16. Amjadi Z, Williamson SS (2010) Power-electronics-based solutions for plug-in hybrid electric vehicle energy storage and management systems. IEEE Trans Ind Electron 57(2):608–616
17. Song Z, Hofmann H, Li J, Han X, Ouyang M (2015) Optimization for a hybrid energy storage system in electric vehicles using dynamic programing approach. Appl Energy 139:151–162
18. Hadartz M, Julander M (2008) Battery-supercapacitor energy storage. Department of Energy and Environment, Chalmers University of Technology
19. Carter R, Cruden A, Hall PJ (2012) Optimizing for efficiency or battery life in a battery/supercapacitor electric vehicle. IEEE Trans Veh Technol 61(4):1526–1533
20. Hannan MA, Azidin FA, Mohamed A (2012) Multi-sources model and control algorithm of an energy management system for light electric vehicles. Energy Convers Manag 62:123–130
21. Allegre AL, Bouscayrol A, Trigui R (2009) Influence of control strategies on battery/supercapacitor hybrid energy storage systems for traction applications. In: IEEE vehicle power and propulsion conference, pp 213–220
22. Pei L, Lu R, Zhu C (2013) Relaxation model of the open-circuit voltage for state-of-charge estimation in lithium-ion batteries. Electr Syst Transp 3:112–117

23. Guo L, Hu C, Li G (2015) The SOC estimation of battery based on the method of improved Ampere-hour and Kalman filter. In: IEEE 10th conference on industrial electronics and applications (ICIEA)
24. He H, Xiong R, Peng J (2016) Real-time estimation of battery state-of-charge with unscented Kalman filter and RTOS lCOS-II platform. Appl Energy 151:1410–1418
25. Sarah K (2008) Spurgeon: sliding mode observers: a survey. Int J Syst Sci 39(8):751–764
26. Yi X, Saif M (2001) Sliding mode observer for nonlinear uncertain systems. IEEE Trans Autom Control 46(12):2012–2017
27. Veneri O, Capasso C, Patalano S (2018) Experimental investigation into the effectiveness of a super-capacitor based hybrid energy storage system for urban commercial vehicles. Appl Energy 227:312–323
28. Song Z, Hou J, Hofmann H, Li J, Ouyang M (2017) Sliding-mode and Lyapunov function-based control for battery/supercapacitor hybrid energy storage system used in electric vehicles. Energy 122:601–612
29. Wang B, Xu J, Xu D, Yan Z (2017) Implementation of an estimator-based adaptive sliding mode control strategy for a boost converter based battery/supercapacitor hybrid energy storage system in electric vehicles. Energy Convers Manag 151:562–572
30. Jung H, Wang H, Hu T (2014) Control design for robust tracking and smooth transition in power systems with battery/supercapacitor hybrid energy storage devices. J Power Sources 267:566–575
31. Emelyanov SV, Korovin SK, Levant A (1996) High-order sliding modes in control systems. Comput Math Model 7:294–318
32. Kamal S, Ramesh Kumar P, Chalanga A, Goyal JK, Bandyopadhyay B, Fridman L (2019) A new class of uniform continuous higher order sliding mode controllers. J Dyn Syst Meas Control. DS-18-1187
33. Hannan MA, Lipu MSH, Hussain A, Mohamed A (2017) A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: challenges and recommendations. Renew Sustain Energy Rev 78:834–854
34. How DN, Hannan MA, Lipu MH, Ker PJ (2019) State of charge estimation for lithium-ion batteries using model-based and data-driven methods: a review. IEEE Access 7:136116–136136
35. Ceraolo M, Lutzemberger G, Poli D (2017) State-of-charge evaluation of supercapacitors. J Energy Storage 11:211–218

Analysis of Reversible Data Hiding Techniques for Digital Images

G. R. Yogish Naik, Namitha R. Shetty, and K. B. Vidyasagar

Abstract With the rapid advancement of information technology, more data and images are now accessible through the Internet. As a consequence, some sort of authentication is required for such sensitive information. Intruders could be present in the transmission path between the transmitter and the receiver and capture the image/data. After the images have been taken, the intruder may be able to view its important content. In other cases, it may not be a problem. Access to such images or data, on the other hand, is not desired in industries such as medicine or the military. Data concealment is one of the most effective methods for avoiding this scenario. Data concealing is a technique in which secret data is integrated into the cover media and is accessible only to the sender and receiver. The method of data concealment has a wide range of applications, including copyright protection for films, videos, and other media. The data concealing method is more secure because attackers cannot alter the information concealed within the cover medium, even if any changes made by the intruder are visible to the sender and recipient. Steganography is perhaps the most effective data concealing methods. One of the uses of steganography is reversible data hide (RDH). After the embedded message has been removed, RDH is a method for retrieving the original cover without losing any data. The following is a synopsis of the research for this subject. A literature review is a brief overview of various techniques proposed by various researchers. The results show that the PSNR for the Lena image is 55.8 dB at a 0.05 bpp embedding rate, which is higher than the TBRDH technique. In terms of PSNR, the suggested LSWTRDH outperforms other embedding rates

Keywords Prediction error · Region of interest · HS histogram shifting · Variable length code

G. R. Yogish Naik (✉) · N. R. Shetty · K. B. Vidyasagar
Department of Computer Science, Kuvempu University, Shivamogga, Karnataka, India
e-mail: ynaik.ku@gmail.com

1 Introduction

Images have risen to become one of the most effective modes of communication. From social media to national security, profile photographs to astronomical imagery, natural beauty to complicated weather patterns, images may convey significant emotion and information more effectively and efficiently than words in a variety of contexts and situations. Modern high-quality digital images, particularly in the fields of space exploration and medicine, have made a significant difference in study and diagnosis.

Images are often employed in covert communication and have shown to be very efficient at concealing sensitive and confidential information. Images include a significant amount of redundant space, which may be efficiently used for further data concealment. Several cryptographic and steganographic techniques are being investigated to see whether they can insert secret and private data into cover images using this feature. With significant technical advancements in hardware, it is now possible to take high-quality images at a cheap cost; the immense expansion of cloud systems, as well as the widespread availability of the Internet, ensures that the use of digital images for communication will continue to grow. One of the most significant and effective uses of steganography is reversible data hiding (RDH). It's the branch that's in charge of secret communication. The key used to encrypt and decode the communication is called the Stego-key. The Stego images are made up of the cover image and the embedded message. Data hiding is a technique for secret communication that uses a digital source as a cover medium and embeds the concealed information into it to make a stego medium [1]. Covert communication, data integrity, temper detection, and copyright protection are just a few of the applications for data concealing techniques. The essential parameters for data concealing techniques are as follows: Payload imperceptibility (or) undetectable security (or) robustness. It's the greatest quantity of concealed information can be stored in the cover media. Bits per pixel are the unit of measurement for embedding capability (bpp). The fundamental concept of data concealing is that it may be done in two ways: one is by embedding the data, and the other is by removing it. In the first instance, covert information is introduced into the cover media during the embedding step. The cover media will be altered as a result of this. Marked/stego data refers to the implanted covert information that has been changed to the cover media.

2 Literature Survey

Wang et al. [2] developed an ideal LSB substitution and genetic algorithm (GA) approaches. Comparing the PSNR for GA to that of simple LSB replacement for Lena image, the PSNR for GA is 44.53 dB, with an average processing time of 0.58 s, which is relatively quick when contrasted to that of less complex LSB substitution for Lena images. Using the Data Encryption Standard (DES) method, Chang

et al. [3] presented a strategy for embedding data that surpassed the Wang et al. approach in terms of security and computation time. Thien and Lin [4] were the first to propose the use of modulus operations for data concealment. In terms of image quality, the results show that modulus-based data hiding surpasses both basic LSB and GA-improved LSB methods. When used to the Optimal Pixel Adjustment Process (OPAP), a simple LSB substitution-based data hiding technique was proposed by Chan and Cheng [5]. The OPAP was used to improve image quality by 51.15 decibels for $k = 1$ while needing minimal computational effort. Yang [6] created an inverted pattern (IP)-based data concealment system that uses testing revealed that every single cover image had a combined average bit capacity of 8, 25,782 bits and a PSNR of 39.11 dB, according to the results method based on exploiting modification direction (EMD) was developed by Zhang and Wang [7], in which each pixel is either raised or lowered by one for total of $(2n + 1)$ unique modifications for the cover pixel is implemented. Zhang et al. [8] devised a double-layered embedding technique that has high embedding efficiency while also having low embedding rates [9–11]. They achieved an accuracy of 50.21, 50.60, 50.45, 50.66, 51.12. The RDH technique for encrypted images developed by Xiaotian Wu and Wei Sun in 2014 [2] is a new RDH approach that makes use of prediction error. Two RDH techniques are given in this suggested system, each of which incorporates the concept of prediction error. A combined method and a separable method are the two distinct types of techniques. As the name implies, in the joint approach, data recovery and image recovery are carried out at the same time, while in the separable technique, data recovery and images recovery are carried out independently of one another. Both methods produce high-quality images with a high embedding capacity and excellent images quality.

3 Proposed Method

With the rapid growth of digital media, ensuring safety and security for web-based communication has become a difficult job. Reversible Data Hiding (RDH) is the basis of steganography and has sparked renewed interest in images in the contemporary day [8, 12, 13]. It hides text, images, and audio/video formats inside another text, image, and audio/video file. RDH has become a fascinating area of image processing as a result of this. Spatial domain and transform domain techniques are two types of reversible data concealing algorithms. Least significant bit (LSB) substitution is the most popular data concealing technique. Researchers divided spatial domain techniques into four categories. Difference expansion is in group I, histogram modification is in group II, and prediction and interpolation algorithms are in groups III and IV, respectively. To be effective, reversible data hiding must follow the triangle rule [14–18], which states that the embedding capacity, or payload, should be large, with greater imperceptibility and security. Security is a major issue among all the other aspects of RDH, and it may be addressed via encryption. It is the process of transforming a images into an unreadable format that safeguards against harmful assaults.

3.1 Secure Force Algorithm

Providing security for digital material has become a difficult job in the contemporary age. Steganography is one such area where data may be reversibly encoded in digital material. The current steganography techniques offer security by encrypting images, concealing data inside it, and then decrypting the image. The issue with Zhang's Ex-OR operation for encryption is that since hider only uses one encryption key, attackers may be able to identify the concealed data. To address this issue, this chapter introduces the Secure Force Algorithm (SFA), which employs several keys for encryption and is therefore immune to malicious assaults. The secure force algorithm is divided into three sections: (a) Key expansion, (b) Encryption, and (c) Decryption. The encrypted images are broken up into mn non-overlapping pieces. The data embedding is then performed in two methods. Depending on the circumstance, the third LSB in the odd rows of blocks is flipped. The final LSB in the even rows of blocks is swapped with the following even block pixels.

(a) **Key expansion block**
Five distinct keys are produced in this phase, each of which performs a different logical and mathematical operation, including permutation and transposition for all of the keys.

(b) **Encryption block**
The encryption procedure starts once the keys produced by the expansion block are retrieved. The procedures for encrypting your data are as follows.

    i. Consider the following input image: 512×512 pixels. Take an 8×8 block of sub-image as an example.
   ii. Arrange 64 pixels components into 16 pixel blocks, one for each of the 16 pixels blocks.
  iii. Using the key k1, do the xnor operation on the first 16 pixel bits and then use the swapping method.
  iv. Using the key k1, do an xnor operation on the final 16 pixel bits before using the swapping method.
   v. Repeat steps iii and iv for each additional key until all of the keys have been finished.

(c) **Decryption**
The reverse process for the above encryption block.

3.2 Data Hiding Algorithm

After the encryption is done by using SF algorithm the data hider starts embedding secret bits as follows (Fig. 1):
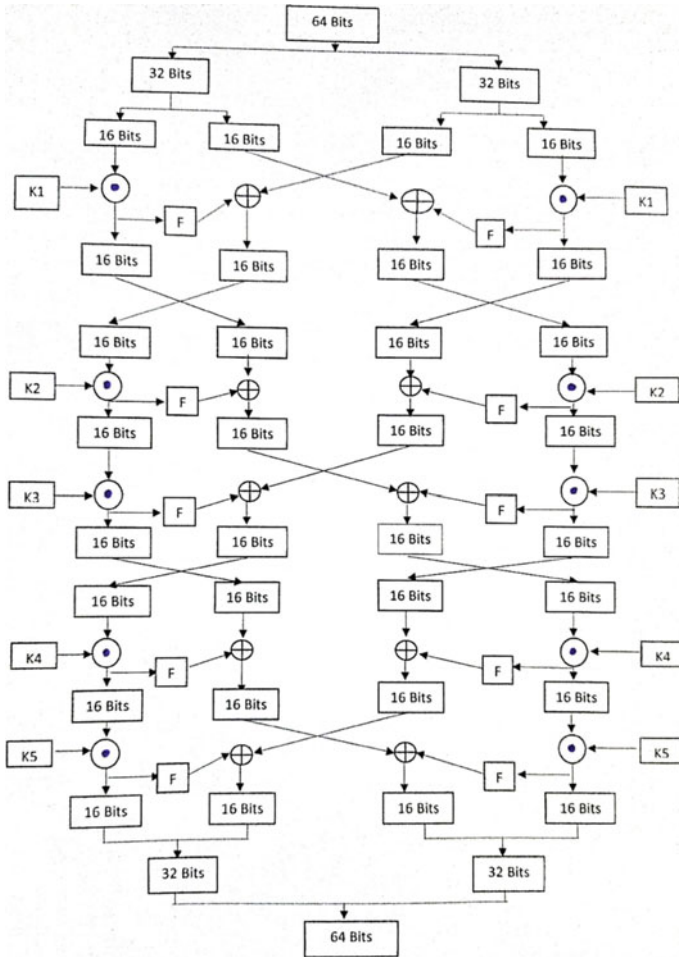
Fig. 1 Bit encryption representation

Step 1: Take a look at the images E of size m × n.

Step 2: The image E is divided into k × k smaller pieces.

Step 3: Using the concealing key, divide each block's 2 k pixels as 0 k and 1 k at random.

Step 4: If binary 0 is to be put, move pixels in 0 k three places. "1" m n q I I 0 (,) m n k and q = 0, 1, 2 (6.5)

Step 5: Change If binary 1 is to be put, move pixels in 1 k three places. "1" m n q I I 1 (,) m n k and q = 0, 1, 2.

Fig. 2 Cover images for simulation

3.3 Extraction Algorithm

Once the secret data has been hidden, the following procedures must be taken to retrieve it: Step 1: Exclusive-OR the decrypted image D and the random sequence generator [19–22]. Step 2: If the inserted bit is 0 and corresponds to pixel 0 k, the three LSB locations are decoded Step 3: If the inserted bit is 1 and corresponds to pixel 1 k, the bits are decrypted in the same manner as in step 2. Step 4: Repeat steps 2 and 3 until all of the bits have been retrieved.

4 Result and Discussion

The simulation results are shown for 512×512 cover images of "Lena," "Airplane," "Barbara," "Baboon," "Boat," and "Peppers" in Fig. 2.

Two parameters are evaluated in this procedure: the first and the second are the SSIM for similarity between the cover images and the rebuilt image. PSNR may be computed using the equation, while SSIM can be determined using the equation. It shows the results of the PSNR and SSIM measurements at various bit rates for the DDH-DCT technique and the SF method. Observing the findings, it can be seen that the PSNR for Boat is 60.25 dB at 0.05 bpp, which is higher than the PSNR for the other cover images, which has an SSIM of 1.0000. The similarity between the cover images and the reconstructed image is measured by the SSIM [23]. When the resemblance is between 0 and 1, the human visual system is unable to differentiate between them. Based on the simulation findings, it can be concluded that the SSIM for the proposed approach is in the range of 0.9955–1.0000. The findings are compared to those obtained using the DDH-DCT technique. By encrypting the cover images using the Ex-OR operation and a single encryption key, this technique makes advantage of data embedding. In order to enhance security, the SF algorithm is described, in which several secret keys are utilized for encryption, followed by data concealment, which protects against hostile assaults. It can be seen from the findings that the proposed SF technique produced excellent visual quality while also having a higher embedding capacity. For the cover images "Lena," higher visual quality of 59.17 dB is obtained at a bit rate of 0.05 bpp while maintaining a PSNR of
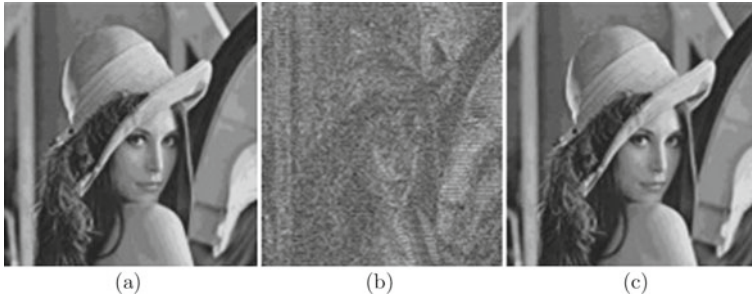
Fig. 3 **a** Cover image **b** encrypted form **c** decrypted form for Lena image

45.45 dB at a bit rate of 0.5 bpp is accomplished. When compared to the DDH-DCT technique, the SF method outperforms it on all five test images with the exception of the baboon, which will be further investigated in the future research. With the exception of baboon, which has an average PSNR of 42.46 dB at various bit rates, all the cover images in the SF technique have higher than 50.07 dB. This is because baboon includes more complex characteristics. A difficult problem is the embedding of data into complex-based images, which is a difficult job. In order to overcome the poor imperceptibility of complex-based images data concealing, it is necessary to develop better methodologies in the future.

The benefit of this technique is that it is more secure since it uses five distinct keys for encryption, making it impossible for attackers to intrude on the data.

5   Conclusion

A new LSB-based RDH model for encrypted digital images, as well as an extra EPR embedding capability, is proposed in this research. Security is improved via the use of encryption, as well as three distinct data embedding mechanisms that provide authentication and extra security in various ways. The three-layer embedding method was developed primarily for the purpose of securing the process and strengthening the authentication measure used throughout the process. The PSNR and SSIM values have both decreased somewhat; nevertheless, the efficiency is still superior to that of the previous techniques in use. Furthermore, the embedding capacity as well as the security is improved. The LSB swapping and substitution procedures are used to carry out the data embedding process. The experimental findings clearly demonstrate that the images quality of the immediately decrypted image is superior to that of the recovered image, and that the image quality is even better in the recovered image. Following data extraction, the error rate is significantly decreased. The capacity for EPR embedding has been enhanced by more than 100% as compared to the prior techniques. All current techniques are significantly outperformed by the suggested

method in terms of PSNR, SSIM, and MSE values as well as embedding capacity, among other metrics.

References

1. Artz D (2001) Digital steganographic: hiding data within data. IEEE Internet Comput 5:75–80
2. Wang RZ, Lin CF, Lin JC (2001) Image hiding by optimal LSB substitution and genetic algorithm. Pattern Recognit 34:671–683
3. Chang CC, Lin MH, Hu YC (2002) A fast and secure image hiding scheme based on LSB substitution. Int J Pattern Recognit Artif Intell 16(4):399–416
4. Thien CC, Lin JC (2003) A simple and high-hiding capacity method for hiding digit-by-digit data in images based on modulus function. Pattern Recognit 36(12):2875–2881
5. Chan CK, Cheng LM (2004) Hiding data in images by simple LSB substitution. Pattern Recognit 37(3):469–474
6. Yang CH (2008) Inverted pattern approach to improve image quality of information hiding by LSB substitution. Pattern Recognit 41(8):2674–2683
7. Zhang X, Wang S (2006) Efficient steganographic embedding by exploiting modification direction. IEEE Commun Lett 10(11):781–783
8. Zhang X, Zhang W, Wang S (2007) Efficient double-layered steganographic embedding. Electron Lett 43(8):482–483
9. Liao X, Wen QY, Zhang J (2011) A steganographic method for digital images with four-pixel differencing and modified LSB substitution. J Vis Commun Image Represent 22(1):1–8
10. Swain G (2014) Digital image steganography using nine-pixel differencing and modified LSB substitution. Indian J Sci Technol 7(9):1444–1450
11. Mielikainen J (2006) LSB matching revisited. IEEE Signal Process Lett 13(5):285–287
12. Chan CS, Chang CC (2007) A survey of information hiding schemes for digital images. IJCSES Int J Comput Sci Eng Syst 1(3):187–200
13. Chang CC, Chou YC (2008) High capacity data hiding for grayscale images. Fundam Inf 86(4):435–446
14. Lee CF, Chang CC, Wang KH (2008) An improvement of EMD embedding method for large payloads by pixel segmentation strategy. Image Vis Comput 26(12):1670–1676
15. Chan CS (2009) On using LSB matching function for data hiding in pixels. Fundam Inf 96(1–2):49–59
16. Li X, Yang B, Cheng D, Zeng T (2009) A generalization of LSB matching. IEEE Signal Process Lett 16(2):69–72
17. Jung KH, Yoo KY (2009) Improved exploiting modification direction method by modulus operation. Int J Signal Process Image Process Pattern 2(1):79–88
18. Yang H, Sun X, Sun G (2009) A high-capacity image data hiding scheme using adaptive LSB substitution. Radio Eng 18(4):509–516
19. Wang ZH, Kieu TD, Chang CC, Li MC (2010) A novel information concealing method based on exploiting modification direction. J Inf Hiding Multimed Signal Process 1(1):1–9
20. Zhang J, Zhang D (2010) Detection of LSB matching steganography in decompressed images. IEEE Signal Process Lett 17(2):141–144
21. Lee CF, Chen HL (2010) A novel data hiding scheme based on modulus function. J Syst Softw 83(5):832–843
22. Luo W, Huang F, Huang J (2010) Edge adaptive image steganography based on LSB matching revisited. IEEE Trans Inf Forens Secur 5(2):201–214
23. Pevny T, Bas P, Fridrich J (2010) Steganalysis by subtractive pixel adjacency matrix. IEEE Trans Inf Forens Secur 5(2):215–224

A Hybrid Cryptosystem of Vigenère and Hill Cipher Using Enhanced Vigenère Table

Nishtha Verma and Ritu Agarwal

Abstract Data security is a huge challenge in this new technical era where everything is handled online. With the increase in population, the size of data is also increasing sharply on a regular basis. Now to protect our data, cryptography plays a major role, Vigenère cipher, which has been used in various platforms for encrypting purposes. Traditional Vigenère cipher uses a table of 26 * 26; however, in this paper, an enhanced and hybrid cryptosystem is proposed, where in place of 26, 62 * 62 table is used, and Hill cipher is combined with it to make it more secure.

Keywords Encryption · Decryption · Security · Cryptosystem

1 Introduction

Size of data is increasing like anything; near about 1.7 megabytes of fresh data is created every second. Data refers to all kind of facts, numerical and textual values. In this online era, every day we are constructing huge amounts of personal and public data. The pandemic scenario, online data usage has risen enormously as more and more people using cloud storage while using Internet banking, adding our details on online sites thus adding data into its databases, online communication in any social network platform that is both textual and multimedia data used for transmission of information, storage of data on cloud. Data breaches or cyberattacks are a major threat to any organization as it can cause a huge loss of revenue to the organization. Data confidentiality and protection of personal data is a major challenge in information technology. Data security helps the organizations to set new and effective digital policies and to prevent unknown access to systems, databases, and various useful sites. Now to give strength to our data and protect it from third parties, cryptography comes to the rescue. Cryptography not only provides information security it also deals with developing some mathematical techniques which can be used to reach goals of cryptography such as—Data confidentiality, data integrity, entity authentication, and

---

N. Verma · R. Agarwal (✉)
Delhi Technological University, Delhi, India
e-mail: ritu.jeea@gmail.com

data origin authentication. With the advancement in technology and hacks, various online fraud and crimes take place and to stop such frauds and crimes, we need strong cipher techniques. Encryption is observed as one strong weapon to keep transmission of data safe over the communication networks.

Process of Encryption and Decryption of Plain Text

Sender—The one who is sending the message. Plain text/message—It is a plain message which the sender wants to send. Encryption algorithm—It refers to the step-by-step process of making some changes in the original data and hence making it understandable only by authorized users. Cipher text—Plain text/message encrypted by the encryption algorithm. Decryption algorithm—It refers to the step-by-step process of changing the cipher text back to plain text. Receiver—One who receives the message.

Objectives of Cryptography

Data Integrity—It is the overall consistency of data to its overall process that is from reaching from to another end. Authentication—It is the process of confirming the origin and integrity of data. Data Confidentiality—Data protection from unlawful access. Non-repudiation—It refers to the assurance from sender that he is processing the data.

Data encryption has many advantages with reference to data security. Encryption algorithms can be classified as—Symmetric key cryptography and asymmetric key cryptography.

(1) Symmetric key cryptography—In these kinds of algorithms, encryption and decryption are done using similar key. Same key is shared between sender and receiver.
(2) Asymmetric key cryptography—Algorithms of these type use two different keys for encryption and decryption one by sender and other by receiver.

Further, symmetric key algorithms are divided into two parts traditional and modern cryptography; traditional ciphers are further divided into two parts: transposition and substitution.

Monoalphabetic ciphers—as the name suggests mono means one letter in plain text is replaced by one cipher text symbol, regardless of its number of occurrences in the message.

Polyalphabetic cipher—In these kinds of ciphers, one letter in plain letter is replaced by other letter, irrespective of the number of occurrences of that letter, each time the letter is replaced by different alphabets. Hence, it is more secure than monoalphabetic.

2 Vigenère Cipher

A 26 × 26 table of row and column with A to Z each is used in Vigenère cipher. This table is generally called as Vigenère square or Vigenère table as shown in Fig. 1. The primary row of this table has 26 English letters. First row starts with A till Z and for the next row each letter receives a single to the left side in a cyclic way. For instance, when C is shifted to the primary place on the third row, the letter B moves to the top, followed by A on top of B. As the Vigenère cipher is a symmetric key encryption algorithm, same key is used for encrypting and decrypting the plain text. Plain text is added with a key to get cipher text. Take the plain text and repeat key stream until length of plain text is exhausted. Select letters from plain text one by one row wise. Select letters from key stream one on one column wise. Find the intersection of row and column. Check in Vigenère table and get intersection letter from table. Write in final cipher text. Repeat the process for each set of letters from plain and key letters.

2.1 Numeric Nature of Vigenère Cipher

Step 1: English alphabets have 26 letters, assign numeric values to each letter starting from 0 to 25.
Step 2: [encryption] Take numeric value of plain text, add it with numeric value of key, and get module 26 of total sum.

Fig. 1 Original Vigenère cipher table [1]

Step 3: [decryption] Take numeric value of plain text, subtract it with numeric value of key, and get module 26 of total sum.

3 Literature Survey

In [2], a new technique is proposed there by using eight new tables after getting cipher text. Each table has different value of letter, also & is added, so module 27 is taken. This technique again has the limitation as it is also case sensitive. In [3], Polybius cipher uses a Polybius square, a table that allows someone to convert letters into numbers. In this proposed method after getting cipher text from Vigenère cipher, text is converted into numerical values using Polybius table, this technique is less secure, as we get number as cipher also only capital letters is used in this method. In [4], researcher included numerical number along with letters, and a key Ukey is taken so, for encryption sum of numerical value of letters of cipher and key then module 36 is taken after that Ukey is added and for decryption subtraction is done.

4 Proposed Approach

Traditional Vigenère cipher uses one key, but in our proposed approach we have two keys: first key we have used for encrypting the message by using Hill cipher and after that we have applied Vigenère cipher technique to generate the final cipher text, also in place of 26 * 26 table we have use 62 * 62 table, so that along with capital letters we can also use small letters and numbers. In traditional Vigenère technique, the plaintext is considered only capital letters; in our approach we have removed this limitation and also have increased security.

4.1 Assign Numerical Values to All Characters

See Tables 1, 2 and 3.

Table 1 Assign 0–25 to uppercase (A–Z)

A	B	C	D	E	F	G	H	I	J
0	1	2	3	4	5	6	7	8	9
K	L	M	N	O	P	Q	R	S	T
10	11	12	13	14	15	16	17	18	19
U	V	W	X	Y	Z				
20	21	22	23	24	25				

Table 2 Assign 26–51 to lowercase (a–z)

a	b	c	d	e	f	g	h	i	j
26	27	28	29	30	31	32	33	34	35
k	l	m	n	o	p	q	r	s	t
36	37	38	39	40	41	42	43	44	45
u	v	w	x	y	z				
46	47	48	49	50	51				

Table 3 Assign 52–61 to numbers (0–9)

0	1	2	3	4	5	6	7	8	9
52	53	54	55	56	57	58	59	60	61

4.2 Encryption Algorithm

Encrypt the message using traditional Hill cipher encryption technique. Hill cipher can be expressed as

- $C = E(K, P) = P * K \mod 26$
- Here, C is for cipher, P for plain text, and K is key.
  - (C1 C2 C3) = (P1 P2 P3) {K11 K12 K13 mod 26
    - – K21 K22 K23
    - – K31 K32 K33}
  - $C1 = (P1K11 + P2K21 + P3K31) \mod 26$
  - $C2 = (P1K12 + P2K22 + P3K32) \mod 26$
  - $C3 = (P1K13 + P2K23 + P3K33) \mod 26$
- Here, the numeric values of plain text are taken and then plain text matrix is multiplied with the key matrix and mod 26 is done.
- Read the cipher text generated from step 1 and take numeric value of each character.
- Select a Vigenère cipher key to apply Vigenère cipher on cipher text generated in step 1.
- Match the length of Vigenère cipher key and cipher text, if the length of Vigenère cipher key is not equal to cipher text, repeat the Vigenère cipher key.
- Take the numeric value of each character for Vigenère cipher key.
- Add the numeric value of cipher text with numeric value of Vigenère cipher key then take module 62.
- Mathematical expression:
  - $F = (C + V) \mod 62$ is final cipher text, and V is Vigenère cipher key
- Convert the numeric values back to corresponding character to generate final cipher.

4.3 Decryption Algorithm

Decrypt the final cipher using the Vigenère cipher key. Vigenère cipher can be expressed as

- $C = (F - V)$ mod 62 is final cipher text, V is Vigenère cipher key, and C is cipher text.
- Read the final cipher text and take numeric value of each character.
- Take the same Vigenère cipher key to decrypt the final cipher text.
- Match the length of Vigenère cipher key and final cipher text, if the length of Vigenère cipher key is not equal to cipher text, repeat the Vigenère cipher key.
- Take the numeric value of each character for Vigenère cipher key and final cipher text.
- Subtract the numeric value of final cipher text with numeric value of Vigenère cipher key and take module 62.
- Now decrypt the cipher text generated by Vigenère cipher using the Hill cipher key.
- Hill cipher can be expressed as
- $P = D(K, C) = C * K^{-1}$ mod 26
- Decryption requires: $K^{-1} = \{1/\text{Det K}\} * \text{Adj K}$
    - (P1 P2 P3) = (C1 C2 C3) {K11 K12 K13 mod 26
        - K21 K22 K23
        - K31 K32 K33}
    - P1 = (C1K11 + C2K21 + C3K31) mod 26
    - P2 = (C1K12 + C2K22 + C3K32) mod 26
    - P3 = (C1K13 + C2K23 + C3K33) mod 26
- Convert the numeric values back to letters using the tables

5 Results

5.1 Encryption

Step 1: Encrypt plain text "PAY MORE MONEY" using Hill cipher with 3 * 3 key matrix. As the key matrix is 3 * 3, we will encrypt three plain texts at the same time (Table 4).

Plain text—PAY MOR EMO NEY

Table 4 Numerical values of the plain text

P	A	Y	M	O	R	E	M	O	N	E	Y
15	0	24	12	14	17	4	12	14	13	4	24

$$\text{Key} - [17\ 7\ 5$$
$$21\ 8\ 21$$
$$2\ 2\ 19]$$

Encrypting: **PAY**, convert the text into numeric form and multiply it with key matrix

$$(C1\ C2\ C3) = (15\ 0\ 24)\ [17\ 7\ 5$$
$$21\ 8\ 21\ \text{mod}\ 26$$
$$2\ \ 2\ 19]$$
$$(C1\ C2\ C3) = ((15*17+0*21+24*2)(15*7+0*8+24*2)$$
$$(15*5+0*21+24*19))\ \text{mod}\ \ 26$$
$$(C1\ C2\ C3) = (303\ 303\ 531)\ \text{mod}\ \ 26$$
$$= 17\ 17\ 11$$
$$= R\ \ R\ \ L$$

Encrypting: **MOR**, convert the text into numeric form and multiply it with key matrix

$$(C1\ C2\ C3) = (12\ 14\ 17)[17\ 7\ 5$$
$$21\ 8\ 21\ \ \ \text{mod}\ 26$$
$$2\ \ 2\ 19]$$
$$(C1\ C2\ C3) = ((12*17+14*21+17*2)(12*7+14*8+17*2)$$
$$12*5+14*21+17*19)\ \ \ \text{mod}\ 26$$
$$(C1\ C2\ C3) = (532\ 490\ 677)\ \ \ \text{mod}\ 26$$
$$= 12\ 22\ 1$$
$$= M\ \ W\ \ B$$

Encrypting: **EMO**, convert the text into numeric form and multiply it with key matrix

$$(C1\ C2\ C3\ ) = (4\ 12\ 14)\ [17\ 7\ 5$$
$$218\ 21\ \ \text{mod}\ 26$$
$$2\ \ 219]$$
$$(C1\ C2\ C3) = ((4*17+12*21+14*2)\ (4*7+12*8+14*2)$$
$$(4*5+12*21+14*19))\ \text{mod}\ 26$$
$$(C1\ C2\ C3) = (348\ 312\ 538)\ \text{mod}\ 26$$
$$= 10\ 0\ 8$$
$$= K\ \ A\ S$$

Table 5 Cipher text generated from step 1

P	A	Y	M	O	R	E	M	O	N	E	Y
R	R	L	M	W	B	K	A	S	P	D	H

Table 6 Encryption using Vigenère cipher key

Cipher text (P)	R	R	L	M	W	B	K	A	S	P	D	H
Numerical value of cipher text	17	17	11	12	22	1	10	0	8	15	3	7
Vigenère cipher key (K)	M	A	Y	O	M	A	Y	O	M	A	Y	O
Numerical value Vigenère cipher key	12	0	24	14	12	0	24	14	12	0	24	14
Sum (Pi + Ki)	29	17	35	26	34	1	34	14	20	15	27	21
Sum % 62 =	29	17	35	26	34	1	34	14	20	15	27	21
Final cipher text	d	R	j	A	k	B	k	O	U	P	b	V

Encrypting: **NEY**, convert the text into numeric form and multiply it with key matrix (Table 5)

$$(C1\ C2\ C3) = (\ 13\ 4\ 24\ )\ [\ 17\ 7\ 5$$
$$21\ 8\ 21 \quad \mod 26$$
$$2\ \ 2\ 19]$$

$$(C1\ C2\ C3) = (\ 13 * 17 + 4 * 21 + 24 * 2\ \ 13 * 7 + 4 * 8 + 24 * 2\ \ 13 * 5$$
$$+ 4 * 21 + 24 * 19)\ \mod 26$$
$$(C1\ C2\ C3) = (532\ 490\ 677)\ \ \mod 26$$
$$= \ 15\ 3\ 7$$
$$= \ P\ \ D\ H$$

Step 2: Now encrypt the cipher obtained in step 1 by using a Vigenère cipher key.
Cipher text—RRLMWBKASPDH.
Vigenère cipher key—MAYOMAYOMAYO.
Mathematical expression for encryption (Table 6):
Final cipher text = (Cipher text + Vigenère cipher Key) % 62.

5.2 Decryption

Step 1: Decrypt the final cipher obtained using a Vigenère cipher key.
Final cipher text—dRjakBKoUpbV.
Vigenère cipher key—MAYOMAYOMAYO.
Mathematical expression for encryption (Table 7):

Table 7 Decryption using Vigenère cipher key

Final cipher text (F)	d	R	j	a	k	B	K	o	U	p	b	V
Numerical value of final cipher text	29	17	35	26	34	1	34	14	20	15	27	21
Vigenère cipher key (K)	M	A	Y	O	M	A	Y	O	M	A	Y	O
Numerical value of Vigenère cipher key	12	0	24	14	12	0	24	14	12	0	24	14
Sum (Fi − Ki)	17	17	11	12	22	1	10	0	8	15	3	7
Sum % 62 =	17	17	11	12	22	1	10	0	8	15	3	7
Cipher text	R	R	L	M	W	B	K	A	S	P	D	H

Cipher text = (Final Cipher letter − Vigenère cipher key letter) % 62.

Step 2: Now decrypt the cipher text obtained in step 1 by using Hill cipher key. For decryption, we need determinant and Adjoint of Key matrix (K) to get value for K^{-1}

$$K^{-1} = \{1/\text{Det K}\} * \text{Adj K}$$
$$\text{Det K} = [\ 17\ 7\ 5$$
$$21\ 8\ 21 \quad \text{mod } 26$$
$$2\ 2\ 19]$$
$$= -939 \ \text{mod } 26$$
$$= 23$$

$$\text{Adj K} = \ |17\ 7\ 5\,|$$
$$|21\ 8\ 21| \quad \text{mod } 26$$
$$|2\ \ 2\ 19|$$
$$= \ |14\ 25\ 7|$$
$$|7\ \ 1\ \ \ 8| \quad \text{mod } 26$$
$$|6\ \ 0\ \ \ 1|$$

By putting the values of Determinant and adjoint to get (K^{-1}) (Table 8):

$$K^{-1} = \{(1/23)\ *\ [\ 14\ 25\ 7$$
$$7\ 1\ \ \ 8 \qquad \text{mod } 26\}$$
$$6\ 0\ 1]$$
$$= \{23^{-1}\ *\ [14\ 25\ 7$$
$$7\ 1\ \ \ 8 \quad \text{mod } 26\}$$
$$6\ \ 0\ 1]$$
$$= \{17\ *\ [14\ 25\ 7$$
$$7\ 1\ \ \ 8 \quad \text{mod } 26\}$$

Table 8 Numerical values of the cipher text

R	R	L	M	W	B	K	A	S	P	D	H
17	17	11	12	22	1	10	0	8	15	3	7

$$
\begin{array}{r}
6 \quad 0 \quad 1] \\
= \quad [\,4 \quad 9 \quad 15 \\
15\,7\,6 \\
24\,0\,17]
\end{array}
$$

Cipher text—RRL MWB KAS PDH.

Decrypting: **RRL**, convert the text into numeric form and multiply it with K^{-1}

$$
\begin{aligned}
(P1\ P2\ P3) &= (R\ R\ L) * \quad [\,4\,9\,15 \\
&\qquad\qquad\qquad\quad 15\,7\,6 \qquad mod\ \ 26 \\
&\qquad\qquad\qquad\quad 24\,0\,17] \\
&= (17\ 17\ 11) \quad [\,4\,9\,15 \\
&\qquad\qquad\qquad\quad 15\,7\,6 \quad mod\ \ 26 \\
&\qquad\qquad\qquad\quad 24\,0\,17] \\
&= (587\ 442\ 544) \quad mod\ \ 26 \\
&= \ 15\ 0\ 24 \\
&= \ P\ A\ Y
\end{aligned}
$$

Decrypting: **MWB**, convert the text into numeric form and multiply it with K^{-1}

$$
\begin{aligned}
(P1\ P2\ P3) &= (M\ W\ B) * \quad 4\,9\,15 \\
&\qquad\qquad\qquad\quad 15\,7\,6 \qquad mod\ \ 26 \\
&\qquad\qquad\qquad\quad 24\,0\,17 \\
&= (402\ 482\ 329) \quad mod\ \ 26 \\
&= \ 12\ 14\ 17 \\
&= \ M\ O\ R
\end{aligned}
$$

Decrypting: **KAS**, convert the text into numeric form and multiply it with K^{-1}

$$
\begin{aligned}
(P1\ P2\ P3) &= (K\ A\ S) * \quad 4\,9\,15 \\
&\qquad\qquad\qquad\quad 15\,7\,6 \quad mod\ \ 26 \\
&\qquad\qquad\qquad\quad 24\,0\,17 \\
&= (472\ 90\ 456) \quad mod\ \ 26 \\
&= \ 4\ 12\ 14
\end{aligned}
$$

Table 9 Numerical values of plain text

P	A	Y	M	O	R	E	M	O	N	E	Y
15	0	24	12	14	17	4	12	14	13	4	24

$$= \text{E M O}$$

Decrypting: **PDH**, convert the text into numeric form and multiply it with K^{-1} (Table 9)

$$(P1\ P2\ P3) = (P\ D\ H) * \begin{array}{ccc} 4 & 9 & 15 \\ 15 & 76 & \text{mod}\ 26 \\ 24 & 0 & 17 \end{array}$$

$$= (273\ 186\ 362)\quad \text{mod}\ 26$$

$$= 13\ 4\ 24$$

$$= \text{N E Y}$$

6 Conclusion

Vigenère cipher is one of the widely used cryptographic techniques. It is considered as weak technique but we are using two keys through which have make it more secure also in traditional Vigenère cipher 26 * 26 table is used, by using 62 table we had removed the limitation of using only capital letters, with our approach all kind of small letters, as well as numbers can also be used which is making it better than earlier.

References

1. The Vigenère cipher encryption and decryption. https://pages.mtu.edu/~shene/NSF-4/Tutorial/VIG/Vig-Base.html. Accessed 25 Feb 2022
2. Soofi AA, Riaz I, Rasheed U (2015) An enhanced Vigenere cipher for data security. Int J Sci Technol Res 4(8):141–145
3. Vatshayan S, Haidri RA, Verma JK (2020) Design of hybrid cryptography system based on Vigenère cipher and Polybius cipher. In: 2020 international conference on computational performance evaluation (ComPE), pp 848–852. https://doi.org/10.1109/ComPE49325.2020.9199997
4. Gautam D, Agrawal C, Sharma P, Mehta M, Saini P (2018) An enhanced cipher technique using Vigenere and modified Caesar cipher. In: 2018 2nd international conference on trends in electronics and informatics (ICOEI). https://doi.org/10.1109/ICOEI.2018.8553910
5. Babu KR, Kumar SU, Babu AV, Reddy MT (2011) An enhanced cryptographic substitution method for information security, pp 2–7

6. Mandal SK, Deepti A (2016) A cryptosystem based on Vigenere cipher by using multilevel encryption scheme. Int J Comput Sci Inf Technol 7:2096–2099

7. Ahamed BB, Krishnamoorthy M (2020) SMS encryption and decryption using modified Vigenere cipher algorithm. J Oper Res Soc China 1–14. https://doi.org/10.1007/s40305-020-00320-x

8. Quist-Aphetsi K (2020) A hybrid cryptosystem based on Vigenere cipher and columnar transposition cipher [Online]. Available at: https://www.academia.edu/5269670/A_Hybrid_Cryptosystem_Based_On_Vigenere_Cipher_and_Columnar_Transposition_Cipher. Accessed 07 Nov 2020

9. Mohanta BK, Jena D, Panda SS, Sobhanayak S (2019) Blockchain technology: a survey on applications and security privacy challenges. Internet Things 8:100107. https://doi.org/10.1016/j.iot.2019.100107

10. Mohammed A-A, Olaniyan A (2016) Vigenere cipher: trends, review and possible modifications. Int J Comput Appl 135(11):46–50. https://doi.org/10.5120/ijca2016908549

11. Abood OG, Guirguis SK (2018) A survey on cryptography algorithms. Int J Sci Res Publ 8(7). https://doi.org/10.29322/ijsrp.8.7.2018.p7978

Intrusion Detection System Using Ensemble Machine Learning for Digital Infrastructure

Merid Nigussie Tulu, Tulu Tilahun Hailu, and Durga Prasad Sharma

Abstract Building the ultimate training model and algorithm selection depends on deep evaluation, knowledge, and interpretation of assessment criterion that is judicious for given domain problems and dataset on hand. Various machine learning studies are focused on extracting knowledge using a single algorithm modeling approach. There is an intuition of selecting and stacking classifiers that are more inclined to yield incomparable performance on a classification problem. Most recently, researchers and professional practitioners are enlightening the ensemble machine learning approach. This paper introduces stacking ensemble classifier algorithms that are used to support vector classifiers (SVC), multi-level perceptron (MLP), and decision tree classifiers (DTC) as base learners, in which their results are aggregated using logistic regression as meta-learners on KDD'99 dataset. During the training, the individual base learner algorithms are fed with massaged KDD'99 datasets. To neglect insignificant features, a principal component analysis is applied with the intuition to extract low-dimensional features with minimal information loss. The K-fold cross-validation was applied in this experiment. The training dataset contains projected dimensions from 41 to 22 dimensions of principal component 96.92%, in which the cutouts data is less significant and accounts for 3.078%. Further, significant and viable efforts are considered for feature engineering, cross-validation, and modeling. After multiple tuned experiments on the model variables, with the performance metrics, a mean and variances are aggregated from 10 iterations. It

M. N. Tulu · T. T. Hailu
Artificial Intelligence and Robotics Center of Excellence, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia
e-mail: merid.nigussie@aastu.edu.et

T. T. Hailu
e-mail: tulu.tilahun@aastu.edu.et

Department of Software Engineering, College of Electrical and Mechanical Engineering, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia

D. P. Sharma (✉)
AMUIT MOEFDRE under UNDP, MAISM RTU, Kota, India
e-mail: dp.shiv08@gmail.com

was done for all the base learners. Finally, the meta-learner result yielded promising concluding remarks.

1 Introduction

Cloud computing is getting popularity, whereas evolving security threats is the main important constraints faced by different scale of cloud service providers and clients [1, 2]. Most recently, computer emergency response team (CERT) report demonstrated the scale of computing security breaches is increasing [2]. Today's big challenges are the dynamic behavior of threats, dataset availability, and domain expertise knowledge. These have led to the idea of developing intelligent systems capable of providing improved intrusion detection. An intrusion detection system (IDS) is an application to detect network threats to protect a computer network from unauthorized access. The main objective of IDS is to give notice to the dedicated system if it suspects the request based on the signature pattern it detects in comparison with existing profile. An IDS process is not a capability of finding a sudden incident in a single moment rather it needs continuous evaluation for a healthy state of the network. Thus, for improved security, the analysis of network logs file is becoming a wide security research area [3].

Machine learning and data mining are two different approaches redundantly used for intrusion detection system. However, most surprisingly, a lot of methodologies significantly overlap, in which machine learning is about forecasting or predicting based on known properties learnt from the training dataset, whereas data mining is about the discovery of unknown properties or knowledge on the given dataset [4]. The designed intrusion detection system model trains and evaluates using KDD'99 dataset that comprises various network traffic attributes.

The process of building intrusion detector system is to build a predictive model (i.e., a classifier) that could be able of distinguishing between bad connections, called intrusions or attacks [5]. Liao et al. [6] emphasized on the dynamic behavior of attacks and security threats as well cited the study of IDSs have received a lot of attention throughout the computer science research area. According to Getachew [7], evaluating the health of the network infrastructure is the way of investigating the security loopholes, which is helpful to enact ICT policies and customized governing standards. Thus, for improved security, the analysis of network logs file is becoming a wide security research area besides the labeling demands domain experts. Suthaharan and Panchagnula [8] argue that the main challenge of current IDS needs continuous data labeling into normal and attacks. Different works classified intrusion detection methodologies into two major categories [9, 10] signature-based detection (SD) and anomaly-based detection (AD) [6, 7] approaches.

1.1 Signature-Based Detection IDS

This approach applies set of rules that describe threats signatures in which attack scenarios are translated into sequences of audit events. The common approach for misuse detection concerns with identified signature verification, whereas a system detects previously seen profiles. The limitation of this approach is the frequent false-alarm rate, and it needs to specify a signature of the attack, and then to update signature of attacks on IDS tools. New attack signatures are not automatically discovered without an update of the IDS signature [11].

1.2 Anomaly Detection-Based IDS

Detection model refers to detection performed by detecting changes in the patterns of behavior of the system. It can be used to detect both known and unknown attack. It can be identified by cataloging and identifying the differences observed. It has the advantage of detecting new attacks over the misuse detection technique. The problem of anomaly-based detection is false positive and false negative reports [1].

Various scholars have been focused to extract knowledge using single algorithm modeling approach. On another hand, stacking ensemble machine learning put on improved mathematical and statistical models together and tools which increase the accuracy of modeling. The main goal of using stacking ensemble-oriented machine learning approach is to detect attacks lies in analyzed, normalized, and transformed KDD'99 dataset and extract knowledge from the dataset for IDS. Thus, this research is to propose an effective ensemble machine learning approach for IDS which focuses on increasing the accuracy of detection rate.

2 Training Dataset

For research purpose, it is difficult to get public dataset for IDSs modeling, whereas experiments have been made on simulated datasets. Bhuyan et al. [12] discussed even though generating new dataset is time-consuming and additional task, they proposed requirement such as real world dataset, completeness in labeling, correctness in labeling, sufficient in trace size for normal and abnormality, concrete feature extraction, diverse attack scenarios, ratio between normal and attack dataset are requirements proposed to generate dataset.

Accordingly, the evaluation of IDS using biased dataset may not fit the solution; however, it is possible to fit the ratio between attack and normal traffic from a large dataset. Having large dataset helps to overcome the time consumed generating new dataset according to the requirement, whereas extracting the subset of the dataset from existing based on the requirements can be used for IDS model evaluation [13, 14].

Table 1 Major categories of attacks

Category	Frequency	Share in percent (%)
DoS	390,479	72.24
Probe	4107	0.83
U2R	52	0.01
R2L	1126	0.23
Normal	97,278	19.69
Total	494,021	100

2.1 Intrusion Detection and Real-Time Log Analysis

Kiatwonghong and Maneewongvatana [15] proposed a real-time log analyzer system that collects and reports cleaned log data from devices and applications in the network and sends to the central server by removing garbage data and defines the relationship among them by converting them into one common format. Learning algorithms such as association rule, term frequency–inverse document frequency (tf-idf), *k*-means, and decision tree were used.

KDD'99 dataset is a public dataset used for many years for IDSs experiment using various techniques [13, 14]. Massachusetts Institute of Technology (MIT) Lincoln Labs [5] labeled each connection as either normal or as an attack. Our work focuses on feature engineering on dataset and applying ensemble machine learning approach to detect the attack type based on the features in KDD'99. The KDD'99 dataset contains four main categories of attacks (i.e., DoS, Probing, R2L, and U2R) with 41 features and one labeled class determines whether the pattern is an attack or normal. Most researchers applied a single algorithm technique to detect all four major attack categories. It is arguable with applying a single algorithm can detect all attacks due to unique execution patterns. The study presented here will create detection models using a comprehensive set of machine learning algorithms applied to attack the categories outlined in Table 1.

3 Data Preprocessing

The original KDD dataset contains a raw TCP dump header logs [16, 17] accordingly, applying proper feature scaling is a requirement. In order to necessitate the differences, some standardization methods have been valued and applied such as feature selection, discretizing, normalization, co-variance matrix, eigenvectors, eigenvalues, and PCA.

4 Stacking Classifiers for IDS

Stacking is an ensemble algorithm where a new model is assembled to combine the predictions from two or more footing models already trained on the dataset [18]. This technique consists of a set of underpinning classifiers algorithms whose individual classification results are combined together to classify new meta-features as a result of their respective predictions. The main intuition behind stacked ensemble learners is that groups of weak learners can come together to form a strong outstanding learner.

Until 2007, the mathematical background of stacking was considered as "Black art," however, Van der Laan et al. [19] introduced the super learner concept, in which they proved that the super learner ensemble represents an asymptotically optimal system for learning.

5 Base Learners

Base learners are always generated from training dataset by using dedicated hypotheses which can be labeled as nominal or continuous value. As a result of the training model, we observed that the ratio is reported, which is converted into a percentage giving an accuracy score of approximately 92.1%. This assures the study argument that the stacked multiple algorithms yield better results than a single algorithm modeling approach. To get the combined better performance, the applied base learners are decision tree, multilayer perceptron, and support vector classifier.

The DT has an easily interpretable structure and is less prone to the curse of dimensionality, which occurs when analyzing and organizing high-dimensional data and DT diagnosing the type of traffic is possible. Multilayer perceptron (MLP) utilizes a supervised learning approach called back propagation for training the network [20, 21]. As in [20], neural network is a powerful algorithm for multi-class classification. Lina et al. [21] combined the advantage of support vector machines (SVMs) and decision tree (DT) and simulated annealing (SA). Lina et al. [21] further proposed the methodologies of a combination of SVM and SA for feature selection and SVM and DT for classification.

6 Meta-Classifier (Meta-Learner)

The base learners' classification models are trained based on the complete training dataset, while the meta-classifier is fitted based on the outputs new meta-features of the individual classification models in the ensemble and specifically logistic regression algorithm used for stacking. Proposed model ensembles the best advantages of DT, SVC, and MLP classifiers, in which their results are aggregated using logistic

regression to boost the performance. A new training dataset constructed from the predictions of the base learners.

```
Algorithm 1: Make predictions on sub-models and
construct a new stacked row

Input: raw dataset for base learners

Output: new dataset from base learners

1: Procedure StackedRow(models, predict_list, row):

2: stacked_row = list()

3:    for i in range(len(models)) do:

4:         prediction = predict_list[i] (models[i],
row)

5:              stacked_row.append(prediction)

6:              stacked_row.append(row[−1])

7:    end for

8: return stacked_row
```

7 Modeling and Training Pipeline

The proposed model is based on individual base learner accuracy. All algorithms discussed in topics DT, MLP, SVC, logistic regression configured with appropriate parameters. The experiment dataset contains projected dimension from 41 to 22 dimensions, in which we used the new reduced dimension $494{,}021 \times 22$ matrix and each attack type is included subsequently K-fold cross-validation with 10 folds cross-validation applied. Ensemble-based machine learning training pipeline has been shown in Fig. 1.

8 Analysis and Principal Component Analysis (PCA) Result

As a result of 22 PCs from information distribution in the experiment and from explained information variance Fig. 2 depicts the loose and information gain. Considering only the front 22 significant PCs with 96.92%, the dimension of the dataset is reduced to $494{,}021 \times 22$ that leads the model training with new reduced dimension.
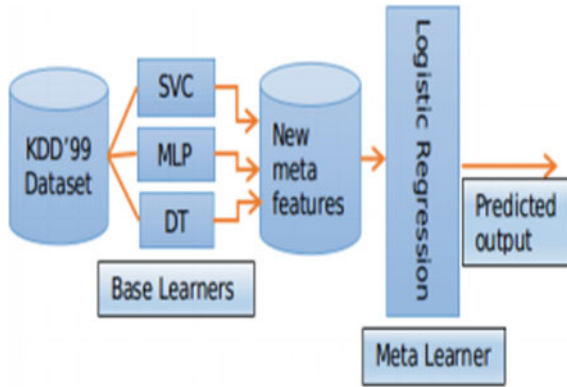
Fig. 1 Ensemble-based machine learning training pipeline
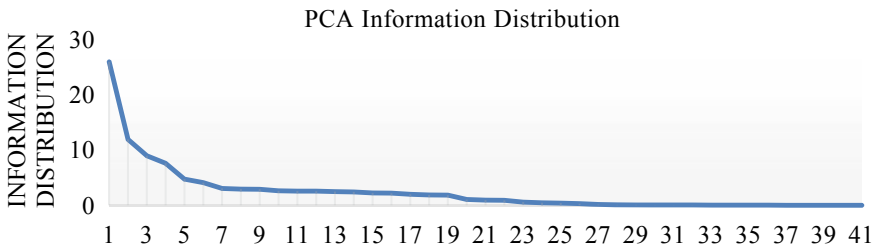
Fig. 2 Principal component analysis result

9 Training and *K*-Fold Cross-Validation

In the context of this study, supervised learning approach applied for classification. The goal is to configure a model that predicts the class of a target variable by learning simple decision rules inferred from the dataset features. To avoid over fitting due to large number of features and to handle discrimination, we applied feature selection and PCA. The training and validation pseudocode are summarized as the following:

```
Algorithm 2: Stacking Classifier with K-Fold [22]
```

Input: Training data for $d = (x_i, y_i)_{i=1}^{m}$ where $(x_i \in R, y_i \in Y)$

Output: Ensemble Stacked Classifiers H

1: Adopt CV technique in preparing a training set for
second-level

2: Split dataset D into equal sized k folds d = {$d1$,
$d2$, …, dk}

3: for $k \leftarrow 1$ to k do

4:     step 1.1: train base classifiers

5:     for $t \leftarrow 1$ to T do

6:     apply a hypothesis h_{kt} on each dataset subset

7:     end for

8:     step 1.2: use the output of base classifier for
second level

9:     for $x_i \in dk$ do

10:         get data (x_i', y^i), where $x_i' = ..., h_{kT}(x_i)$}

11:     end for

12:     end for

13: Step 2: Train the second level classifier

14: Train a new classifier on the new data set {x_i, y_i}

15: Repeat training base classifiers

16: for $t \leftarrow 1$ to T do

17: apply a hypothesis on dataset

18: end for

19: return $H(x) = h'(h1(x), (h1(x), \ldots (hT(x))$

Totally, taking the 10-folds and use ninefold CV to get out-of-sample predictions for each of the first layer models on all 10-folds. The leave out one of 10-folds and fit the models on the other nine and then predict on the held-out data. The process repeats for all nine folds so we get out-of-sample predictions on all nine folds (Fig. 3). Considering the cost of experiment, tenfold cross-validation is realistic.

Confidence Interval (CI). Besides the accuracy as evaluation metric, we introduced statistical inference to estimate a confidence interval. It gives a likely range of values and the estimated range being calculated from a given dataset.
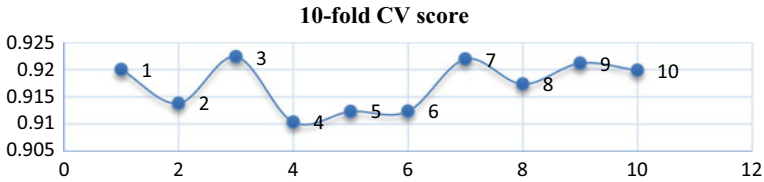
Fig. 3 Result of model tenfold CV

Table 2 Model prediction and result

Algorithm	Accuracy	MSE	CI
[MLP]	90.8 Std: (+/ − 0.098)	− 0.050 (0.150)	[0.997425, 0.999163]
[SVC]	91.8 Std: (+/ − 0.100)	− 0.100 (0.199)	[0.997171, 0.998946]
[DT]	91.7 Std: (+/ − 0.134)	− 0.150 (0.229)	[0.996560, 0.998930]
[Stacking]	92.1 Std: (+/ − 0.178)	− 0.050 (0.150)	[0.989648, 0.992794]

$$\text{Confidence interval} : \left(x - z^* \frac{\sigma}{\sqrt{n}}, x + z^* \frac{\sigma}{\sqrt{n}} \right), \tag{1}$$

where x is the mean of the score, σ is the standard deviation, and z^* is the constant critical value chosen based on the confidence level on the score reported, for this study, 2.26 [23] is taken considering 99%. The prediction and results of different algorithms are shown in Table 2.

10 Conclusion

The significance of the used techniques is noticeably focused on data preprocessing and the implementation of the stacked ensemble method. In the first place, the study conjugated the prior gap of data distribution viable in the train/test approach that K-fold cross-validation applied. Secondly, the study considered stacked learners to benefit from the ensemble techniques using multi-class classifiers machine learning algorithms on the original dataset. This was done for both training and evaluation. For final prediction, a meta-classifier logistic regression is used as a stacking algorithm considering the probability error resilience.

Based on the results revealed and analyzed, it was concluded that the KDD'99 dataset information is in a way that allows dealing with intrusion detection in network traffic with a higher and advanced detection rate of accuracy. Hence, the overall performance (92.1%) of the proposed model is determined to have an improvement in detection accuracy. In conclusion, the paper provides an enhanced and promising direction for the construction of better IDS based on a stacked ensemble machine learning approach.

References

1. US-CERT federal incident notification guidelines | US-CERT. https://www.us-cert.gov/incident-notification-guidelines. Accessed 03 Mar 2017
2. Wu SX, Banzhaf W (2010) The use of computational intelligence in intrusion detection systems: a review. Appl Soft Comput J 10:1–35
3. Lu CT, Boedihardjo AP, Manalwar P (2005) Exploiting efficient data mining techniques to enhance intrusion detection systems. 0-7803-9093-8/05/$20.00. IEEE, pp 512–517
4. Jabez J, Muthukumar B (2015) Intrusion detection system (IDS): anomaly detection using outlier detection approach. Procedia Comput Sci 48:338–346
5. KDD-CUP-99 task description. http://kdd.ics.uci.edu/databases/kddcup99/task.html. Accessed 02 Mar 2017
6. Liao HJ, Richard Lin CH, Lin YC, Tung KY (2013) Intrusion detection system: a comprehensive review. J Netw Comput Appl 36(1):16–24
7. Getachew S (2015) Layer based log analysis for enhancing security of enterprise data center. Doctoral dissertation, Addis Ababa University
8. Suthaharan S, Panchagnula T (2012) Relevance feature selection with data cleaning for intrusion detection system. In: 2012 Proceedings of IEEE South-eastcon
9. Hossain M (2004) Intrusion detection with artificial neural network. M.Sc. thesis, Technical University of Denmark, Lyngby
10. Yan KQ, Wang SC, Liu CW (2009) A hybrid intrusion detection system of cluster-based wireless sensor networks. In: Proceedings of the international multiconference of engineers and computer scientists, vol 1
11. Planquart JP GSEC Certification-version 1.2d 'Application of neural networks to intrusion detection'
12. Bhuyan MH, Bhattacharyya DK, Kalita JK (2015) Towards generating real-life dataset for network intrusion detection
13. Olusola AA, Oladele AS, Abosede DO (2010) Analysis of KDD '99 intrusion detection dataset for selection of relevance features. In: World Congress on Engineering and Computer Science. International Association of Engineers, pp 162–168
14. The 1998 intrusion detection off-line evaluation plan. MIT Lincoln Lab., Information Systems Technology Group. http://www.11.mit.edu/IST/ideval/docs/1998/id98-eval-11.txt
15. Kiatwonghong N, Maneewongvatana S (2010) Intelli-log: a real-time log analyzer. In: 2nd International conference on education technology and computer (ICETC), pp 22–24
16. Manandhar P (2014) A practical approach to anomaly-based intrusion detection system by outlier mining in network traffic. Acknowledgments-semantic scholar
17. Aggarwal CC, Yu PS (2001) Outlier detection for high dimensional data. In: Proceedings of the 2001 ACM SIGMOD international conference on management of data. New York, pp 37–46
18. Zhou ZH (2015) Ensemble Learning. In: Li SZ, Jain AK (eds) Encyclopedia of biometrics. Springer US, pp 411–416
19. Van der Laan MJ, Polley EC, Hubbard AE (2007) Super learner. U.C. Berkeley division of biostatistics working paper series. Working paper 222
20. Cunningham R, Lippmann R (2000) Improving intrusion detection performance using keyword selection and neural networks. R-MIT Lincoln University
21. Lina S-W, Yingb K-C, Leec C-Y, Leed Z-J (2012) An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection. Appl Soft Comput 12(10):3285–3332
22. Tang J, Alelyani S, Liu H (2015) Data classification: algorithms and applications, Data mining and knowledge discovery series. CRC Press, pp 498–500
23. Confidence intervals. http://www.stat.yale.edu/Courses/1997-98/101/confint.htm. Accessed 04 Jun 2017

Integrating InceptionResNetv2 Model and Machine Learning Classifiers for Food Texture Classification

Philomina Simon and V. Uma

Abstract Modern agricultural revolution has enabled farmers to use Artificial Intelligence (AI)-based technologies resulting in improved crop yield, reduction in cost and optimization of the agricultural processes. Most of the food crops and grains are cultivated through farming and each food has its own characteristics, texture, color, shape, size, and granularity. Recognizing and identifying food items is necessary for calorie and nutrition estimation in order to maintain healthy eating habits and good health. Since the food items come in various textures, contents, and structure, it is tedious to distinguish food material. Texture is a prominent feature usually observed in all the agricultural food grains, and it is difficult to classify food material without recognizing its texture. It is evident that texture analysis and recognition is a significant and vital component in agriculture-based applications. The intention of the work is to prove the significance of the deep architecture InceptionResNetv2 model integrated with machine learning classifiers and to assess the performance of the classifiers in food classification system. We tested our proposed model with challenging RawFooT food texture dataset with 68 classes using10 fold cross validation. Extensive ablation study and analysis demonstrated an excellent accuracy of 100% with linear discriminant analysis and 99.8% with Quadratic SVM classifier for RawFooT dataset.

Keywords Food texture · InceptionResNet · Deep feature · Quadratic SVM · Linear discriminant analysis · Classification

P. Simon (✉)
Department of Computer Science, University of Kerala, Thiruvananthapuram, Kerala 695581, India
e-mail: philominasimon@keralauniversity.ac.in

V. Uma
Department of Computer Science, Pondicherry University, Puducherry, India
e-mail: uma.csc@pondiuni.edu.in

1   Introduction

Due to the advent of the Artificial Intelligence (AI)-based intelligent systems, the agriculture has shifted its focus to precision farming. Majority of the food that we consume are cultivated though agriculture. In this modernized agriculture revolution era, there are deep learning based techniques that can be adopted for increasing the agricultural productivity [1]. Today agriculture remains the most important economic activity in the world. Precision agriculture is a term used for AI-enabled computer aided approach to farming management to analyze the needs of the individual crops. Recognizing food materials can help human beings control the eating habits and stay healthy. Texture is a significant feature that helps to identify the feel, structure and roughness of an object usually present in the food grain and crops. Nowadays, agricultural image processing is an upcoming research area which helps in the automation of different tasks involved in crop classification, food classification and leaf disease classification. These methods employ deep learning techniques for performing classification and efficient decision making. In this work, we investigate the effectiveness of texture classification for food grain classification. Texture is an inevitable feature descriptor involved in the agricultural image processing. Deep learning extends the machine learning techniques by automatically extracting the prominent features from the given data. The highlight of the work is applying AI-based computer vision techniques for food texture classification, investigating the efficiency of deep features with machine learning classifiers such as SVM, K-NN, Naïve Bayes, random forest, and identifying the best ML classifier for food texture classification. In this work, we investigate the efficiency of using deep learning model InceptionResNet with the machine learning classifiers. We inferred that InceptionResNet based deep feature generation is found excellent when integrated with Fisher LDA, SVM, and random forest classifiers.

1.1   *Significance of Texture Analysis in Food Recognition*

Texture is a prominent feature that is used in computer vision applications. The food texture by Matz [2] is an earlier work for food classification systems. Texture features are commonly used for agricultural image classification and recognition tasks. Texture features depict the characteristics of the food grains and the food images [3]. Chaugule and Mali [4] investigated the significance of shape and texture classification for paddy varieties. Food texture and structure [5] are significant in analyzing the food materials. The presence of the moisture and fat determines the food texture. Food grains are agricultural particles that are characterized with shape, texture, and size. The performance of the classification algorithm depends on the deep extracted features which can be further processed to identify food grains.

2 Related Work

Shen et al. [6] presented a prototype for AI-based food recognition system using pretrained deep learning models, namely, Inceptionv3 and Inceptionv4 and also performed textual data modeling for estimating the attributes in the food image. Liu et al. [7] in the work applied an extra 1×1 convolution layers to the inception block and applied it in the AlexNet architecture for multi-level food recognition and obtained a Top 5 accuracy of 95.2% in UEC100 dataset. Patil et al. [8] studied the influence of color models and developed a method by merging the color and haralick texture features with K-NN classifier for 10 class food grain classification. Accuracy of 95.93% was obtained. Subhi et al. [9] discussed different computer vision and AI-based techniques for automated food recognition and dietary assessment. Okada et al. [10] captured the food images with a sensor and classified 3 classes of food texture with a recurrent neural network. Pinto et al. [11] elaborated on the machine learning approach, K-means clustering for leaf disease segmentation and classification based on the gray level co-occurrence matrix (GLCM), texture and color features in sunflower plants. Cusano et al. [12] developed the RawFooT database and demonstrated that CNN descriptors outperformed the handcrafted features in food texture classification.

3 Deep Learning Approach for Food Texture Classification by Proposed Method

The proposed model illustrated in Fig. 1 extracts the deep features using Inception-ResNetv2 and performs classification using machine learning classifiers. Inception-Resnet is developed after being inspired by the advantages of skip connections in ResNet models and the inception blocks. In InceptionResNet deep model, the filter concatenation of the inception block is replaced with the residual connections. This model take advantage of the ResNet and the inception models there by retaining the computational efficiency [13].

3.1 Automated Feature Extraction Using Deep Networks

The greatest benefit of the deep learning lies in the automated feature learning from high level to low level features in the texture feature space. Deep Learning is widely used in plant disease detection, crop type classification, and food texture classification [14]. In this work, we use InceptionResNet [13], pretrained CNN for deep feature extraction. The InceptionResNet block is depicted in Fig. 2.
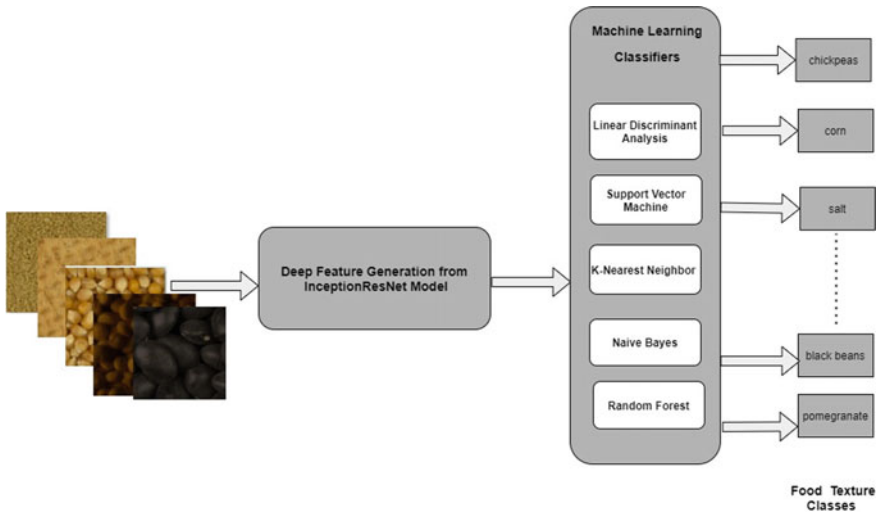
Fig. 1 Deep learning for food texture classification by proposed method

With the influence of pretrained models, deep learning became efficient in terms of computational simplicity and feature learning. Skip connections or residual connections are implemented by Hu et al. [15]. The significant benefit achieved by using a transfer learning model is extraction of deep image features through the pretrained models rather than training a new CNN. This results in simplicity and computational efficiency. CNN process images only once and extracts the deep feature representations. Better feature representations and accuracy are thus obtained from limited data. In convolutional neural network, the convolution filters derive prominent key features in the image and pretrained CNN layers produce responses (features) of the food image. These 'activations' or deep feature representations can be automatically extracted in the form of a feature vector from each convolution layers ranging from edges, corners, regions and textures.

We have used InceptionResNetv2 pretrained model for extracting the features. We adopted a better approach for deep feature extraction rather than developing a custom CNN, and we have used these deep features to train machine learning classifiers. The pretrained CNN we have chosen achieved high accuracy results when compared with the state-of-art-methods. Training or learning texture features from the scratch within a CNN requires the image data to be organized as mini batches and it is iteratively sent through many conv layers, pooling or subsampling stages. The weight for CNN needs to be learned from back propagation algorithm to compute the stochastic gradient descent to minimize the error rate.

To train the network, it requires enormous computational power and we need to tune the hyperparameters properly. The method of deep feature extraction requires less computational power. For extracting significant texture features from deep
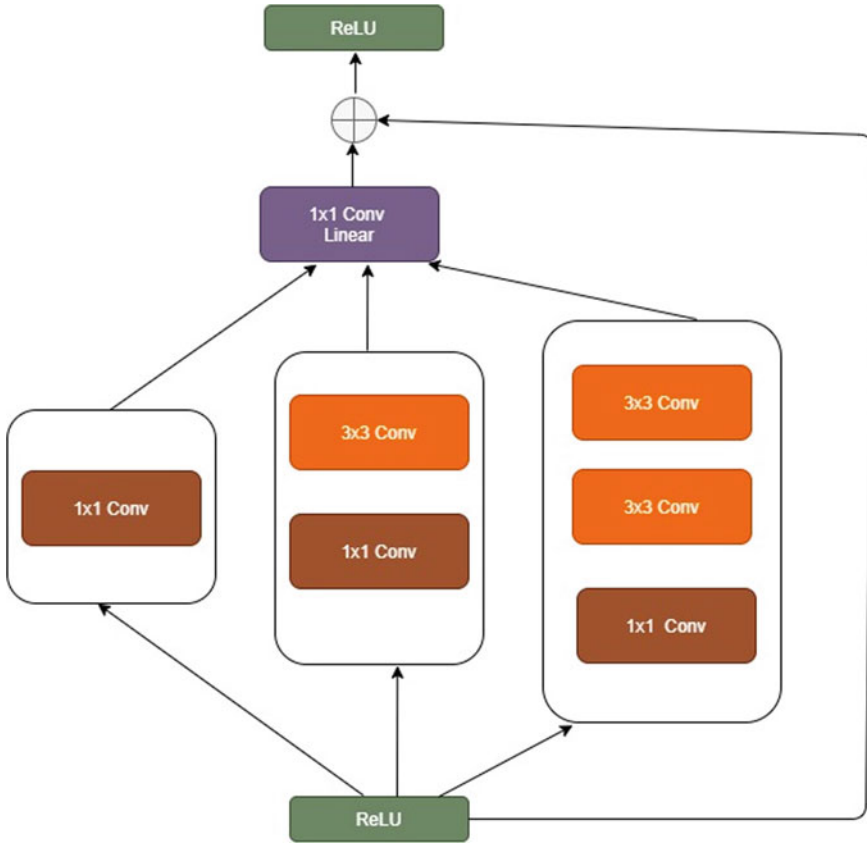
Fig. 2 InceptionResNet block

networks, less number of epochs is sufficient. Pretrained CNNs is a straightforward and less complex approach where we can perform deep feature extraction, classification, and transfer learning in the presence of limited training images more effectively.

3.2 Machine Learning Classifiers

The significance of Fisher Linear Discriminant Analysis (LDA) [16] lies in the fact that it can be used as a feature reducer and classifier. In this food classification approach, LDA is used as a classifier to estimate the optimal features that separates the classes based on intra-inter class variances. Support vector machine (SVM) is a supervised model that classifies the features in the high-dimensional texture space [17]. Polynomial kernel of order 2, quadratic kernel is used in SVM and one vs

all strategy is used for multiclass classification. Quadratic SVM can separate the nonlinear texture features in a better way. K-Nearest Neighbor (K-NN) algorithm is a lazy learner or an instance based algorithm where the K neighborhood points are assigned to the majority class. The value of K is set as 3 and the distance measure used is euclidean distance. Naïve Bayes classifier is based on Bayes theorem that uses conditional probability. Random Forest [18] is an ensemble-based decision tree classifier. This tree bagging classifier uses the parameter, number of tree learners (set as 30) and learning rate (set as 0.1) for the model.

4 Experimental Analysis and Discussion

4.1 RawFooT Food Texture Dataset

RawFooT [5, 19] food texture image data classify food image textures under various illumination scenarios. Images in the dataset are captured based on the intensity variations, color, illumination, and light direction. The dataset comprises of 68 classes and each class comprises of 92 images [12]. RawFooT dataset contains food types ranging from fruit, cereals, meat, etc. In the work, we investigated the significance of InceptionResNet deep features for food classification.

4.2 Results

Cross validation is a good technique for evaluating the deep learning models and it gives the information how a classifier generalizes the residuals of the classifier. In this work, we performed tenfold cross validation scheme for evaluating the performance of the classifier. In tenfold cross validation, the dataset is split into 10 subgroups; i.e., S_1, S_2, S_3, S_4………S_{10}, where each subgroup is of same size. From these 10 groups, one group is considered as test set and the remaining 9 groups perform as training set. The process is iterated 10 times until all the group is assigned as the test set and train set. The results of ten subgroups are combined, and average accuracy is calculated. The trained model is made less biased and optimized. We tested our model in Raw Food Texture database (RawFooT) which illustrates the texture of the food grains with 46 conditions of illumination, color, orientation, and intensity with five machine learning classifiers.

The deep features are extracted from InceptionResNetv2 model and classification with LDA, Quadratic SVM, K-NN, Naïve Bayes, and Random Forest. Our work concluded that best accuracy is obtained with 100% accuracy for Fisher LDA, 99.8% for Quadratic SVM and 99.7% for tree bagging algorithm Random Forest. The results are shown in Table 1. Mcallister et al. [20] analyzed the significance of applying deep learning architectures such as GoogleNet and ResNet152 for recognizing the variety

of food with artificial neural network classifier and obtained 99% and 99.2% accuracy. Cusano et al. applied VGG-19 fully connected network as a transfer learning model and got 98.21% results. Figure 3 represents the receiver operating characteristics (ROC) curve for food classification with InceptionResNetv2 and LDA classifier.

Table 1 Evaluation results for food classification with InceptionResNet features and ML classifiers

Methods	CNN Model	Classifier	Accuracy (%)
Proposed method (Best result 1)	InceptionResNetv2	Linear discriminant analysis	100
Proposed method (Best result 2)	InceptionResNetv2	Quadratic support vector machine	99.80
Proposed method (Best result 3)	InceptionResNetv2	Random Forest	99.70
Proposed method	InceptionResNetv2	K-NN	97.80
Proposed method	InceptionResNetv2	Naïve Bayes	96.80
Cusano et al. [12]	VGG19	CNN-FC	98.21
Mcallister et al. [20]	ResNet152	Artificial neural network	99.20
Mcallister et al. [20]	GoogLeNet	Artificial neural network	99.00
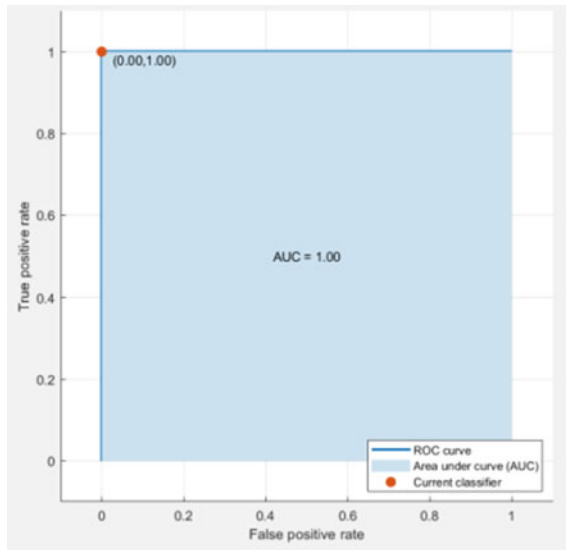
Fig. 3 Evaluation results for food classification with InceptionResNet and LDA classifier

5   Conclusion

AI-based intelligent technologies aid to significant improvements in the agricultural revolution. Productivity in the crop and the food grain cultivation can be improved by the usage of the technological advancements such as deep learning and texture. In this work, we present a deep learning model InceptionResNetv2 with machine learning classifiers for efficient food texture image classification. Experimental analysis demonstrates the superiority of the proposed method when compared with other methods. We obtained accuracy of 100, 99.8, and 99.7% using Fisher LDA, Quadratic SVM, and Random Forest, respectively, when tested on RawFooT food texture dataset.

References

1. Tyagi AC (2016) Towards a second green revolution. Irrig Drain 4(65):388–389
2. Matz SA (1962) Food texture. Avi Publishing Company
3. Koga T et al. (2003) Characterization and classification of foods by texture of food. J Jpn Soc Food Sci Technol
4. Chaugule A, Mali SN (2014) Evaluation of texture and shape features for classification of four paddy varieties. J Eng
5. Rosenthal A, Chen J (2015) Modifying food texture, volume 1: novel ingredients and processing techniques. Woodhead Publishing Series in Food Science, Technology and Nutrition. https://doi.org/10.1016/C2014-0-02669-X
6. Shen Z et al. (2020) Machine learning based approach on food recognition and nutrition estimation. Procedia Comput Sci 174:448–453
7. Liu C et al. (2017) A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure. IEEE Trans Serv Comput 11(2):249–261
8. Patil NK, Malemath VS, Yadahalli RM (2011) Color and texture based identification and classification of food grains using different color models and haralick features. Int J Comput Sci Eng 3:3669–3680
9. Subhi MA, Ali SH, Mohammed MA (2019) Vision-based approaches for automatic food recognition and dietary assessment: a survey. IEEE Access 7:35370–35381
10. Nakamoto H, Nishikubo D, Okada S, Kobayashi F, Kojima F (2017) Food texture classification using magnetic sensor and principal component analysis. In: Proceedings of the 2016 3rd international conference on computing measurement control and sensor network, pp 114–117. https://doi.org/10.1109/CMCSN.2016.39
11. Pinto LS, Ray A, Reddy MU, Perumal P, Aishwarya P (2017) Crop disease classification using texture analysis. In: Proceedings of the 2016 IEEE international conference on recent trends in electronics, information and communication technology (RTEICT 2016), pp 825–828. https://doi.org/10.1109/RTEICT.2016.7807942
12. Cusano C, Napoletano P, Schettini R (2016) Evaluating color texture descriptors under large variations of controlled lighting conditions. J Opt Soc Am A 33:17
13. Szegedy C et al. (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-first AAAI conference on artificial intelligence
14. Kamilaris A, Prenafeta-Boldú FX (2018) Deep learning in agriculture: a survey. Comput Electron Agric 147:70–90
15. He K et al. (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition

16. Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann Eugen 7(2):179–188
17. Kim KI et al. (2002) Support vector machines for texture classification. IEEE Trans Pattern Anal Mach Intell 24(11):1542–1550
18. Breiman L (2001) Random forests. Mach Learn 45:5–32
19. RawFooT DB: raw food texture database (2015). http://projects.ivl.disco.unimib.it/rawfoot/
20. McAllister P et al. (2018) Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets. Comput Biol Med 95:217–233

Sentiment Analysis of Customer on a Restaurant Using Review in Twitter

Nagaratna P. Hegde, V. Sireesha, G. P. Hegde, and K. Gnyanee

Abstract As it is important to maintain the customer relationship the business firms like restaurants, Apps, etc., should take the opinion of the users and improve from the feedback they will give to the company and this improves the relationship with the customer promising his/her regular visits to your businesses. In this research, we will see the results of using natural language processing in analyzing the restaurant reviews dataset and by which categorization into positive and negative reviews restaurant owner can save his time by taking only the negative reviews into consideration and can improve their businesses.

Keywords Naive Bayes classifier · VADER · Sentiment analysis · Confusion matrix

1 Introduction

As technology is growing on par with the population, the complexity of analyzing is increasing as a result of huge data that is generated day by day. So, the technologies like natural language processing, machine learning are useful for analysis of this huge data. This research is aimed to improve a restaurant's functioning by analyzing whether the customer had liked the food or not in that restaurant based in the review given by the customer.

N. P. Hegde (✉) · V. Sireesha
Vasavi College of Engineering, Hyderabad, Telangana, India
e-mail: nagaratnaph@staff.vce.ac.in

V. Sireesha
e-mail: v.sireesha@staff.vce.ac.in

G. P. Hegde
SDM Institute of Technology, Ujre, India

K. Gnyanee
NCR, Hyderabad, India
e-mail: gk185089@ncr.com

2 Proposed Scheme

The proposed scheme analyzes the opinions about restaurant and attempts to understand the sentimentality behind the users on a large scale. The high-level system design of the proposed structure is given in the Fig. 1. This structure is used to quantify opinion estimation of restaurant. The suggested framework mainly consists of dataset, pre-processing, feature extraction and sentiment determination modules [1]. These modules are discussed in detail in the below sections.

2.1 Data Set

The study aims to test user opinion as it changes in real-time. So, reviews that match the subject of the study are extracted from the live timeline and stored as ".tsv" file. The dataset is of size 1000×2 and the columns are review of the customer regarding the restaurant and the label telling whether it is a positive or negative review.

As shown in Fig. 2, the review in the dataset is mapped against the target values either '0' or '1'. '0' indicating it is negative review and '1' representing it is a positive review. Here, "Liked" variable is a target variable, which is used in training the model.

Fig. 1 Proposed scheme
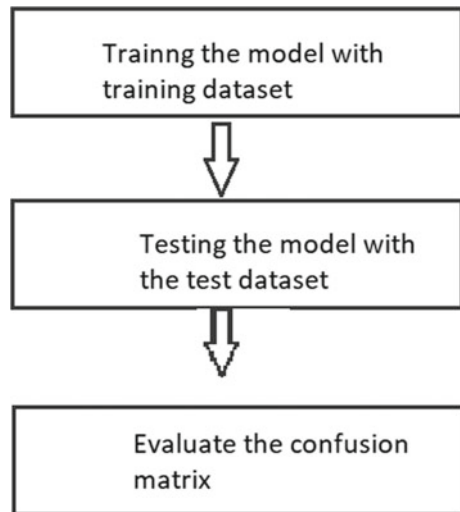
Fig. 2 Restaurant review dataset

	Review	Liked
0	Wow... Loved this place.	1
1	Crust is not good.	0
2	Not tasty and the texture was just nasty.	0
3	Stopped by during the late May bank holiday of...	1
4	The selection on the menu was great and so wer...	1
...	...	...
995	I think food should have flavor and texture an...	0
996	Appetite instantly gone.	0
997	Overall I was not impressed and would not go b...	0
998	The whole experience was underwhelming, and I ...	0
999	Then, as if I hadn't wasted enough of my life ...	0

2.2 Data Pre-processing

Reviews are very unstructured, with grammatical and non-grammatical errors, linguistic variants, and an abundance of acronyms and emoticons. As a result, processing this data is critical before performing any statistical analysis on it. The NLTK library is used to perform data pre-processing. It is a Python package that provides an easy-to-use API for text-processing activities. On the reviews, the NLTK Library is utilized to execute conventional natural language processing (NLP) activities. The raw data is subjected to a succession of data formatting processes:

Original sentence: Wow… Loved this place.

Cleaning: All the exclamations are removed.

```
review = re.sub('[^a-zA-Z]',' ',dataset['Review'][i])
```

o/p: Wow Loved this place.

Tokenization: For further processing, a token is a collection of characters that are put together as a meaningful unit, such as symbols or words. Tokenization is the process of extracting the tokens from the tweet content. Internally, tokens are kept in a database.

```
review=word_tokenize(review)
```

o/p: ['wow', 'loved', 'this', 'place'].

Stemming: It is the process of reducing a derived word to its stem by normalizing text. For example, stemming the phrases "swimmer," "swam," and "swimming" yields the root word "swim." To simplify word comparison, stemming is employed to ignore the complex grammatical transformations of the term in the original text. Potter

Stemmer is employed in this instance. The stop words are then eliminated in this stage.

```
review=[ps.stem(word) for word in review if word not in stop ]
```

o/p: ['wow', 'love', 'place'].

Stop-words removal: A few terms like "a," "an," "the," "he," "she," "by," "on," and others are employed solely for grammatical and referential purposes and have no bearing on data analysis. These terms, known as stop words, can be eliminated because they are common but provide no more information about the topic.

Stop words is a module from the "NLTK" package that is used for this.

```
nltk.download('stopwords')
```

2.3 Feature Extraction

The Features in the processed data is extracted by using countVectorizer function from sklearn.feature_extraction.text module. Here, we are restricting the features by an upper limit of "1500".

```
from sklearn.feature_extraction.text import CountVectorizer
cv=CountVectorizer(max_features=1500)
x=cv.fit_transform(corpus)
```

This gives an array of the features that a particular review contains. For example,

```
corpus=[
... 'This is the first document.',
... 'This document is the second document.',
... 'And this is the third one.',
... 'Is this the first document?',]
```

If countVectorizer is applied to this corpus then the feature array will be like

```
['and', 'document', 'first', 'is', 'one', 'second', 'the', 'third',
'this']
```

These are the Features extracted and Feature array is:

```
[[0 1 1 1 0 0 1 0 1]
 [0 2 0 1 0 1 1 0 1]
 [1 0 0 1 1 0 1 1 1]
 [0 1 1 1 0 0 1 0 1]]
```

2.4 Sentiment Determination

The processed data can be used to use statistical methods to determine the sentiment of a specific review, as well as the sentiment of the entire set of data. Classification and polarity assignment are the two subtasks of this method.

Classification. The content from the reviews is now fed into a computer model, which assigns categories to each tweet representing a review. Classification is the categorizing of data based on how they are similar, i.e., similar patterns occur in one class of items but differ from the patterns in the other. A classifier is constructed in the proposed work to accurately categorize tweets into two sentiment categories: positive and negative.

The system is divided into two parts, as shown in the Fig. 3, training and prediction. The quantitative model is trained using a pre-classified dataset so that it can examine and learn the features that are classified into different classes. The data from the processed restaurant reviews is then fed into the trained model, which extracts features. The features are fed into a classifier model, which uses the underlying mathematical machine learning method to assign polarity. The classifier model is chosen to provide the highest level of accuracy for the type of data being used. In this work, the Naive Bayes classifier is applied and result is compared with the results obtained by Vader tool.

Naive Bayes is a probabilistic machine learning technique that calculates the joint probability of words and specified classes to obtain the probabilities of classes given to texts using Bayes' theorem. The Naive Bayes algorithm predicts the likelihood of each class's membership of a specific word. The most likely class is assigned to the one with the highest probability. The mathematical reasoning behind the calculation of the unigram word models using Naive Bayes is represented by the following equation. The fundamental idea is to calculate the likelihood of a word appearing in any of the potential classes from a given training sample.

$$P(\text{word}/\text{obj}) = P(\text{word in obj})/P(\text{Total words in obj})$$

The likelihood of whether a review is objective is computed using Bayes' rule, which involves estimating the probability of a review from the specified objective class and the earlier probability of the objective class. The term $P(\text{review})$ can be substituted with $P(\text{review} \mid \text{obj}) + P(\text{review} \mid \text{subj})$ [2].

$$P\left(\text{obj}/\text{review}\right) = \frac{P\left(\text{Tweet}/\text{obj}\right)P(\text{obj})}{P(\text{Tweet})}$$

If the unigrams inside the review are assumed to be independent (i.e., the occurrence of one word in the review does not affect the probability of occurrence of any other word in the review), the probability of a review given the objective class can be computed as a simple product of the probabilities of all words in the review belonging to the objective class. Furthermore, if comparable class sizes for objective

Fig. 3 Training and prediction modules
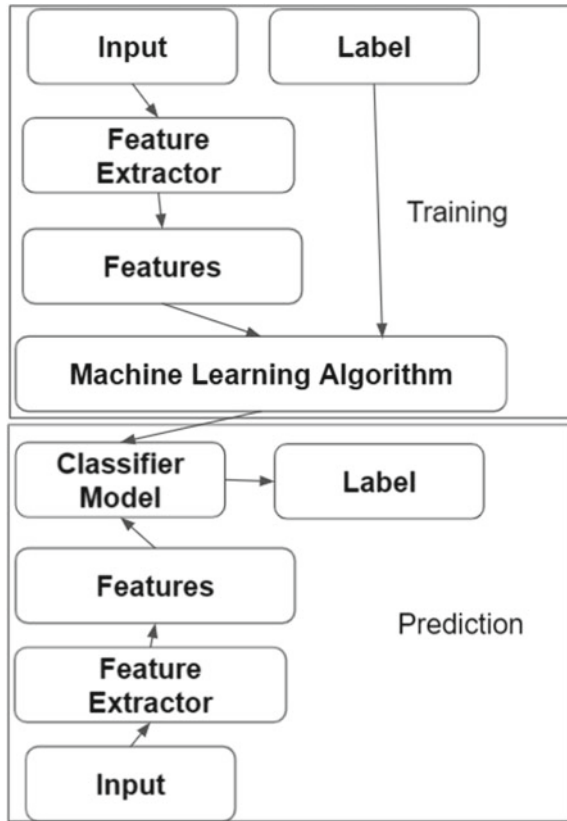
and subjective classes are assumed, the objective class's prior probability can be ignored. The following equation can now be written, with 'wi' standing for the ith word.

$$P(\text{obj/review}) = \frac{\prod_{i=1}^{N} P(\text{wi}/\text{obj})}{\prod_{i=1}^{N} P(\text{wi}/\text{obj}) + \prod_{i=1}^{N} P(\text{wi}/\text{sbj})}$$

The Maximum A Posteriori (MAP) approach is used to determine the highest likelihood of a word's class. The MAP for a hypothesis is determined by the equations below, where $P(H)$ is the hypothesis probability and $P(E)$ is the evidence probability [2].

$$\text{MAP}(H) = \max(P(H|E)) = \max\big((P(E|H) * P(H))\big/ P(E)\big)$$
$$= \max(P(E|H) * P(H))$$

Polarity Assignment. The system's last module focuses on determining if a sentiment is favorable, negative, or neutral. Following the text categorization task, sentiment analysis is performed. Contextual sentiment analysis and broad sentiment analysis are the two types of sentiment classifications. Contextual sentiment analysis is concerned with classifying certain elements of a review based on the context. The sentiment of the entire text is examined in general sentiment analysis. Because the study's purpose is to determine user sentiment regarding the restaurant, the proposed schema follows general sentiment analysis. As a result, it's critical to comprehend the sentiment expressed in the evaluation as a whole. The classifier used in this case, the Naive Bayes classifier, is a probabilistic classifier, meaning that for a document 'd,' out of all classes $c \in C$ the classifier returns the class $\hat{c}$, which has the maximum posterior probability given the document. The following equation denotes the estimate of the appropriate class.

$$c = \arg\max\ P(c \mid d) \text{ where } c \in C$$

The "NLTK" package for Python is used to perform the natural language processing (NLP) tasks [3]. Sentiment analysis is performed on the parsed by using the sentiment polarity method of the TextBlob library. The tweets are assigned a polarity score in the range $[-1.0, 1.0]$.

sentiment.polarity $> 0 \Rightarrow$ '*positive*', sentiment.polarity $== 0 \Rightarrow$ '*negative*'

The model assigns a subjectivity score in the range $[0.0, 1.0]$ in addition to a polarity score. 0.0 is a fairly objective value, while 1.0 is a very subjective value. An intensity score is applied to every word in the lexicon, along with polarity and subjectivity, to determine how much it alters the next word.

Consider the term 'very.' It has a 0.3 polarity score, a 0.3 subjectivity score, and a 1.3 intensity score. By adding 'not' before this word, the polarity is multiplied by -0.5 while the subjectivity remains constant. 'Very' is noted as a modifier during POS tagging. TextBlob ignores polarity and subjectivity in the phrase 'very great,' but considers intensity score when calculating the sentiment of the next word, 'great.' In this way, TextBlob assigns polarity and subjectivity to words and phrases and averages the scores for longer sentences/text.

According to sentiments expressed in the tweets, the tweets are labeled in three classes: positive, negative, and neutral.

Positive: Review is considered as positive if the user expressed review has a love/great/prompt/amazing/cute or if the words mentioned have positive annotations [4].

Negative: Review is considered as negative if the review has negative/angered/sad feelings or if words with negative annotations are mentioned.

3 Results

The study's goal is to understand customers' opinions toward a restaurant using reviews given to restaurant. The reviews are analyzed by the model and a confusion matrix is generated. To test the model built the result is compared with the result obtained by VADER considering the 1000 tweets. Performance of the proposed work based on Naïve Bayes classification is measured by finding accuracy and precision using confusion matrix. Valence Aware Dictionary and sEntiment Reasoner (VADER) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. The classification of tweets as positive and negative by proposed method as well as by VADER is given in Table 1 and its comparison is shown in Fig. 4. Tables 2 and 3 present the confusion matrix of sentiment analysis by proposed method as well as VADER [5].

Accuracy: It may be defined as the number of correct predictions made by our ML model. We can easily calculate it by confusion matrix with the help of following formula [6].

Table 1 Classification details

	Negative tweets	Positive tweets
Dataset classification of tweets	348	652
Tweets classified by VADER	375	625
Tweets classified by proposed method	382	618
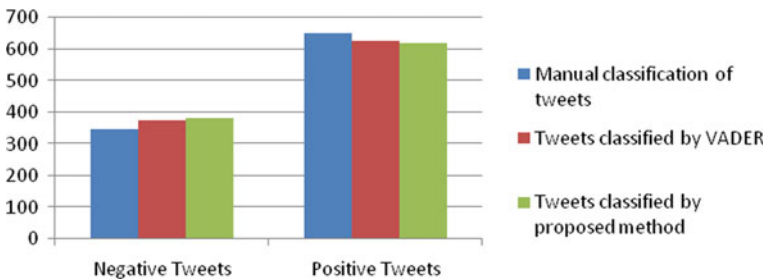
Fig. 4 Comparison of classification results

Table 2 Confusion matrix for the proposed method

Total tweets = 1000		True tweet class	
		Positive	Negative
Predicted tweet class	Positive	610	8
	Negative	42	340

Table 3 Confusion matrix for the sentiment analysis by VADER

Total tweets = 1000		True tweet class	
		Positive	Negative
Predicted tweet class	Positive	600	25
	Negative	52	323

Table 4 Accuracy comparison

Accuracy by Naïve Bayes (proposed) (%)	Accuracy by VADER (%)
95	92

Fig. 5 Accuracy
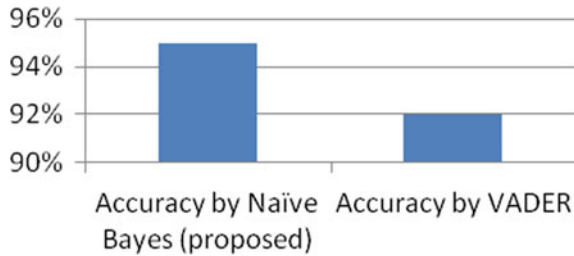
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

A comparison of accuracy obtained by proposed model and by VADER has been shown in Table 4 and Fig. 5.

Precision: It is defined as the number of correct documents returned by our ML model. We can easily calculate it by confusion matrix with the help of following formula [7].

$$\textbf{Precision} = \frac{TP}{TP + FP}$$

The precision of our model is 0.98 and by VADER is 0.96.

4 Conclusion

The results show that the proposed model is effective when compared to VADER. Our findings emphasize that most of the people love the restaurant as the number of the '1's are more in prediction done by the proposed model. So, many people are likely to come to that restaurant as they have given the good reviews.

References

1. Rosenberg H, Syed S, Rezaie S (2020) The twitter pandemic: the critical role of twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. Can J Emerg Med 22(4):418–421
2. Padmavathia S, Ramanujam E (2015) Naïve Bayes classifier for1 ECG abnormalities using multivariate maximal time series motif. Procedia Comput Sci 1(47):222–228
3. Jianqiang Z, Xiaolin G (2017) Comparison research on text pre-processing methods on twitter sentiment analysis. IEEE Access 5:2870–2879
4. Chakraborty AK, Das S, Kolya AK (2021) Sentiment analysis of covid-19 tweets using evolutionary classification-based LSTM model. In: Proceedings of the research and applications in artificial intelligence. Springer, Singapore, pp 75–86
5. Imran AS, Doudpota SM, Kastrati Z, Bhatra R (2020) Cross-cultural polarity and emotion detection using sentiment analysis and deep learning—a case study on COVID-19. arXiv preprint arXiv:2008.10031
6. Kaur H, Ahsaan SU, Alankar B, Chang V (2021) A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets. Inf Syst Front 1–13
7. Al-Subaihin SA, Al-Khalifa SH (2014) A system for sentiment analysis of colloquial Arabic using human computation. Sci World J

Fundus Image Processing for Glaucoma Diagnosis Using Dynamic Support Vector Machine

K. Pranathi, Madhavi Pingili, and B. Mamatha

Abstract In this paper, an efficient fundus image processing (FIP) is presented to diagnose glaucoma using dynamic support vector machine (DSVM). It has three important elements: preprocessing, texture feature extraction, and classification. The Laplace of Gaussian Filtering (LGF) is used for enhancing the high frequency components of fundus images at first. A stochastic texture field model (STFM) is used to extract the textures. Then, the DSVM classifier, an enhanced version of SVM is utilized for the classification. The proposed FIP system is evaluated using RIM-ONE and ORIGA database fundus images with tenfold cross validation. Data augmentation is also employed for increasing the samples. Results show that the FIP system provides 97.2% correct classification on RIM-ONE and 85.5% accuracy on ORIGA database images. Also, it is noted that features extracted after preprocessing have more discriminating power for glaucoma classification than the features from the original fundus image.

Keywords Fundus image · Glaucoma · Dynamic support vector machine · Texture features · Laplace of Gaussian filter

1 Introduction

The cupping of optic nerve head (ONH), visual field loss, and the elevated intraocular pressure are the characteristic of the glaucoma. For better diagnosis, these characteristic should be identified at the earliest. Micro statistical descriptors based system is discussed in [1] for glaucoma diagnosis. It uses neural networks to classify the micro structures obtained using Laws filters. Laws features such as level, ripple, spot, edge, and wave are extracted from the ONH only. Convolution neural network (CNN)-based glaucoma diagnosis system to run in mobile devices is discussed in [2]. It uses fundus images from different databases. Two CNNs are designed for segmentation of ONH and for the classification of glaucoma with confidence level.

K. Pranathi (✉) · M. Pingili · B. Mamatha
Department of IT, CMR Engineering College, Hyderabad, India
e-mail: write2pranathi@gmail.com

A hybrid system is designed in [3] using pre-trained CNN for glaucoma diagnosis. It extracts the ONH using bottom-up visual focus paradigm. Different computations are made to classify the stages of glaucoma. The segmentation of ONH including optic disc and optic cup and the extraction of retinal features such as color and textures are discussed in [4]. Retinal features are then classified using dynamic ensemble classifiers. The ONH presents in the region of interest and background are separated more accurately.

Spectral featured for glaucoma diagnosis is discussed in [5]. Different sampling rates are used in stock well transform to extract low frequency and high frequency components of fundus image. These components are classified using Random Forest (RF) classifier. The color channels in the retinal images are analyzed in [6] for glaucoma diagnosis. ONH can be clearly visible in the red components and the blood vessels in the green channel.

A two subsystems for effective glaucoma diagnosis is described in [7] using CNN. First system utilizes machine learning approaches for ONH segmentation and classification from physical and positional features. Second system utilizes pre-trained CNN for classification then the outputs are combined for final decision. A cup-to-disk ratio (CDR)-based diagnosis system is discussed in [8]. It is a direct method without segmentation and formed as a regression problem by a RF classifier.

ONH segmentation and CDR for analyzing glaucoma is discussed in [9]. K-means and Fuzzy c means are employed for the segmentation and the hill climbing algorithm determines the number of clusters. A global semantically diagnosis model is discussed in [10]. It provides clinically interpretable diagnosis by the use of deep learning. It aggregates features from different scales for better performance.

In this paper, an efficient FIP system is discussed for glaucoma diagnosis. The rest of the paper is organized as follows: Sect. 2 discusses the design of the proposed FIP system by the use of LGF, texture features and DSVM. The performances of the FIP system is analyzed in Sect. 3 using two fundus image databases and Sect. 4 gives the conclusion of the FIP system.

2 Methodology

This section discusses the design of the proposed FIP system using DSVM. It has three important elements; preprocessing by LGF for high frequency enhancement, texture feature extraction by STFM, classification by DSVM. Figure 1 shows the flow of the proposed system for glaucoma diagnosis using DSVM.

2.1 Preprocessing

The LGF enhances the high frequency information of an image (edges, lines on animate) and it is thus a high pass filter. It is defined as [11]

| Fundus Images | → | LGF (Preprocessing) | → | Textures by STFM | → | DSVM (classifier) |

```
Fundus          LGF            Textures by      DSVM
Images    →   (Preprocessing)  →   STFM    →   (classifier)
                                                   │
                                                   ↓
                                              Glaucoma
                                              diagnosis
```
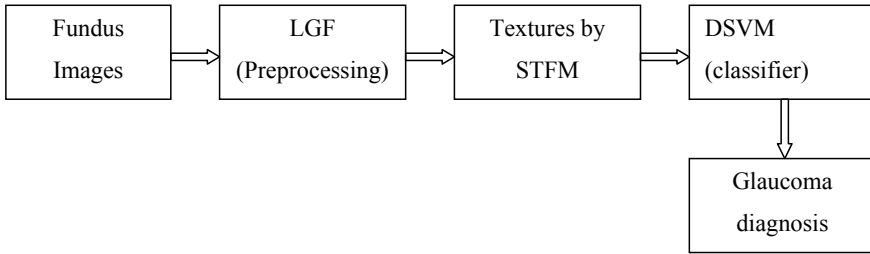
Fig. 1 Flow of FIP system for glaucoma diagnosis

$$\text{LGF}(x, y) = \left(\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \right) e^{-\left(\frac{x^2 + y^2}{2\sigma^2} \right)} \tag{1}$$

where x and y are pixel co-ordinates in an image, and σ is the filter's standard deviation. Its implementation involves the convolution of an image with a Gaussian function. The Gaussian filter's standard deviation determines the amount of smoothing. The next stage involves convolving the Gaussian-smoothed image with a Laplacian kernel to produce the gradient of an image. The Laplace transform could not be used alone because it is a gradient filter which makes it sensitive to noise. The Gaussian was used prior to the use of the Laplacian to suppress noise. However, the LGF is sensitive to noise as it amplifies high frequencies. Pre-convolving the Gaussian with a Laplacian filter to form a hybrid filter is economic and computationally cheaper as there is only one convolution with an image.

A large positive value is normally added to a LGF filter's response to eliminate negative values resulting from the filtering. As σ increases filtering becomes insignificant and the LGF progressively approaches the.

Laplacian kernel. In this work, σ is a smoothing parameter and different σ expose different scales of features? This is the key in computer vision as the information of interest might be easily captured from one scale than the other. An optimization of σ and also the Gaussian window size are therefore critical for the success of filtering. In this work, σ is set to 0.5, and Gaussian window size of 3×3 is used.

2.2 STFM Features

The statistical pattern recognition algorithms consist of several procedures that include the choice of features and the choice of decision or classification strategies. The term features is used mainly to one of the measurements from fundus image samples. The performance of the pattern recognition algorithms is primarily based on how the extracted features are accurate to discriminate the normal and abnormal samples. In this work, STFM is used for the generation of fundus textures. According to this model, an array of identical distributed random variables passes
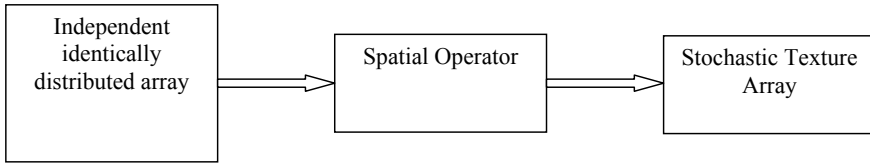
Fig. 2 STFM

through a linear or nonlinear spatial operator to produce a stochastic texture array. Figure 2 shows the STFM.

The statistical properties of the STFM are determined by the generating probability density of the identically distributed array (IDA) and the spatial operator. Therefore, the probability density and the spatial operator provide a complete description of such a texture. The whitening transformation based upon the field's autocorrelation function de-correlates the texture field, and provides an estimate of IDR. However, it is not clear how the spatial operator could be estimated; the problem being made particularly difficult by IDR not being known exactly. It is suggested in [12] that instead of using spatial operator as a second texture descriptor, the autocorrelation function of the texture field denoted as KF should be used as it contains most of the texture information not evident in probability density of IDA. It is conceded that these two descriptors provide an incomplete description of the texture field, but it is suggested that they provide sufficient information for the majority of applications. Obvious features to use are the moments of the estimated probability density function and two dimensional spread measures of the autocorrelation function KF.

2.3 DSVM Classification

The standard SVM [13] is a binary classifier which, when given a test example, outputs the class of the samples by the geometric distance to the optimal separating hyper plane. It is based on SVM classification of linear subspace representations. In this approach, a single SVM is trained to distinguish between the normal and glaucoma images, respectively. The inherent potential of the SVMs in extracting the relevant discriminatory information from the training data, irrespective of representation and pre-processing gives promising results. In conventional SVM, the regularization constant is fixed which reduces the accuracy of the system. To overcome this, the regularization constant is replaced by an exponentially increasing constant in DSVM [14]. The empirical error in conventional SVM is defined as

$$\text{SVM} = C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \tag{2}$$

$$\text{DSVM} = C \times \frac{2}{1 + e^{p1 - \frac{2p1i}{l}}} \sum_{i=1}^{l} (\xi_i + \xi_i^*) \tag{3}$$

where C is the regularization constant

3 Results and Discussions

This section discusses the performances of the FIP system for glaucoma diagnosis. It uses two different databases; RIM-ONE and ORIGA database images. The former one contains only the ONH cropped image and the later one has full fundus image with 3072×2048 pixels. Figure 3 shows the sample fundus images from both databases.

ORIGA database has 168 glaucoma and 482 normal images whereas RIM-ONE database has 39 (glaucoma) and 92 (normal) images. In order to train the classifier properly, the samples are increased using data augmentation. After this process, number of glaucoma images are increased to 273 (RIM-ONE) and 1176 (ORIGA) and normal images to 3374 (ORIGA) and 644 (RIM-ONE). Then tenfold validation

(a)

(b)
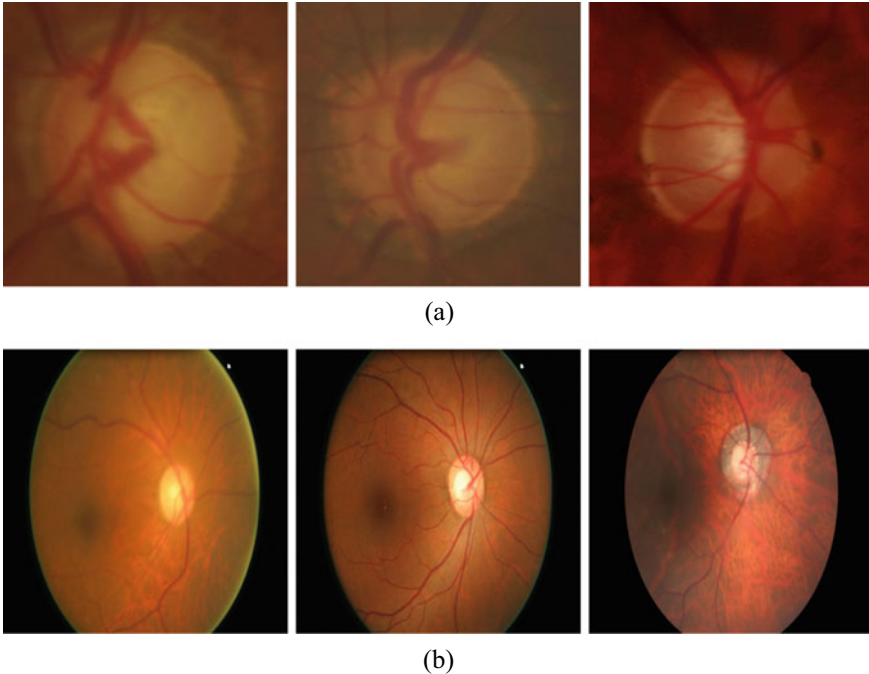
Fig. 3 Sample fundus images, **a** RIM-ONE, **b** ORIGA

is employed for the classification to split the data into training and testing. Figure 4 shows the confusion matrix obtained by the FIP system using DSVM without preprocessing. Figure 5 shows the confusion matrix obtained by the FIP system using DSVM after preprocessing.

It can be seen from the Figs. 4 and 5 that the FIP system with LGF provides better results than without preprocessing for RIM-ONE and ORIGA databases. The FIP system LGF gives 97.2% accuracy and 85.5% accuracy for RIM-ONE and ORIGA databases which is 3% (RIM-ONE) and 5.7% (ORIGA) improvement over FIP system without LGF element in the design. The maximum sensitivity obtained by the FIS system with LGF is 95.6% for RIM-ONE and 84.9% for ORIGA database. Also,
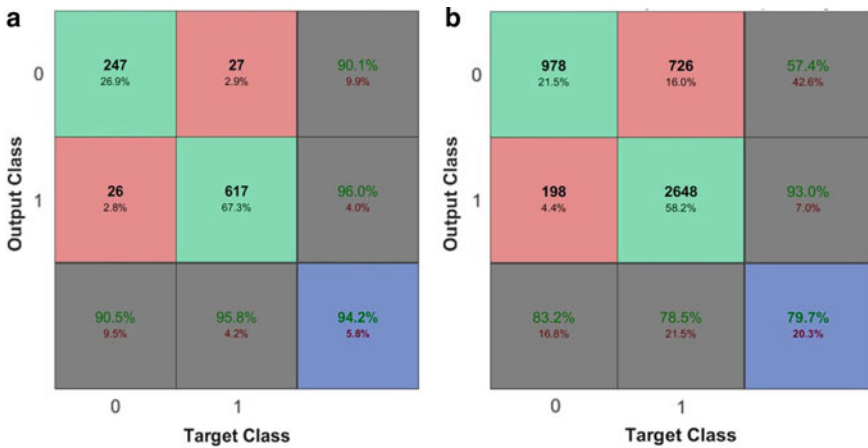
Fig. 4 Confusion matrix of FIP system without preprocessing, **a** RIM-ONE, **b** ORIGA
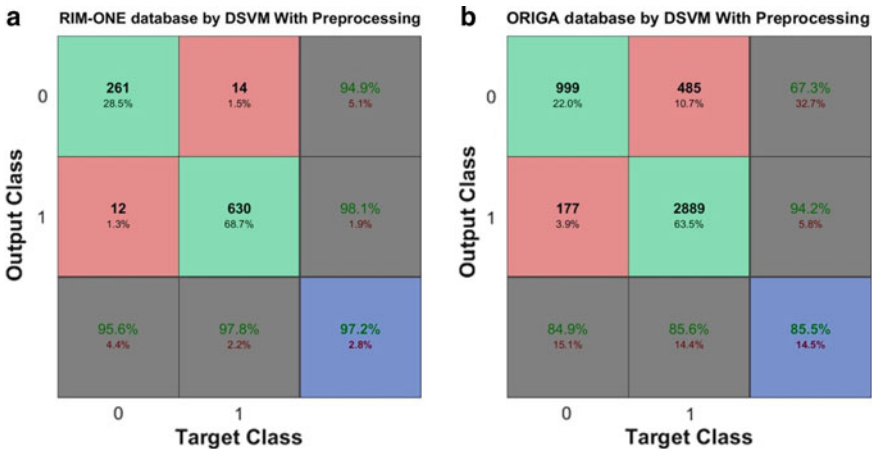
Fig. 5 Confusion matrix of FIP system with LGF, **a** RIM-ONE, **b** ORIGA

the maximum specificity obtained is 97.8% for RIM-ONE and 85.6% for ORIGA database. The performance of the FIP system on ORIGA database can be improved by cropping the ONH as in RIM-ONE database.

4 Conclusion

The proposed FIP system has explored the use of machine learning approach to classify the fundus image samples from texture features. It turns out that the combination of image enhancement and texture features gives a good result. The FIP system includes LGF, texture features by STFM and DSVM. Experiments are conducted on RIM-ONE and ORIGA database fundus images including preprocessing and excluding preprocessing element. The extracted texture features are useful in characterizing the abnormality in the fundus image for glaucoma diagnosis. Results proved that the combination of LGF and texture features provides the best results when classifying using DSVM. At best, the proposed FIP system achieved 97.2% corrected classification on the normal and glaucoma samples of RIM-ONE and 85.5% for ORIGA samples.

References

1. Alagirisamy M (2021) Micro statistical descriptors for glaucoma diagnosis using neural networks. Int J Adv Signal Image Sci 7(1):1–10
2. Martins J, Cardoso JS, Soares F (2020) Offline computer-aided diagnosis for Glaucoma detection using fundus images targeted at mobile devices. Comput Methods Programs Biomed 192:1–13
3. Subha K, Kother Mohideen S (2021) The efficient computer aided diagnosis (CAD) systems for detecting glaucoma in retinal fundus images using hybrid pre-tained CNN. Des Eng 16666–16681
4. Pathan S, Kumar P, Pai RM, Bhandary SV (2021) Automated segmentation and classifcation of retinal features for glaucoma diagnosis. Biomed Signal Process Control 63:102244
5. Gokul Kannan K, Ganeshbabu TR (2017) Glaucoma image classification using discrete orthogonal stockwell transform. Int J Adv Signal Image Sci 3(1):1–6
6. Satya Nugraha G, Amelia Riyandari B, Sutoyo E (2020) RGB channel analysis for glaucoma detection in retinal fundus image. In: 2020 International conference on advancement in data science, e-learning and information systems, pp 1–5
7. Civit-Masot J, Domínguez-Morales MJ, Vicente-Díaz S, Civit A (2020) Dual machine-learning system to aid glaucoma diagnosis using disc and cup feature extraction. IEEE Access 8:127519–127529
8. Zhao R, Chen X, Liu X, Chen Z, Guo F, Li S (2020) Direct cup-to-disc ratio estimation for glaucoma screening via semi-supervised learning. IEEE J Biomed Health Inform 24(4):1104–1113
9. Ganeshbabu TR (2015) Computer aided diagnosis of glaucoma detection using digital fundus image. Int J Adv Signal Image Sci 1(1):1–11
10. Liao W, Zou B, Zhao R, Chen Y, He Z, Zhou M (2020) Clinical interpretable deep learning model for glaucoma diagnosis. IEEE J Biomed Health Inform 24(5):1405–1412

11. Jingbo X, Bo L, Haijun L, Jianxin L (2011) A new method for realizing LOG filter in image edge detection. In: Proceedings of 2011 6th international forum on strategic technology, pp 733–737
12. Faugeras OD, Pratt WK (1980) Decorrelation methods of texture feature extraction. IEEE Trans Pattern Anal Mach Intell 2:323–332
13. Bakare YB, Kumarasamy M (2021) Histopathological image analysis for oral cancer classification by support vector machine. Int J Adv Signal Image Sci 7(2):1–10
14. Cao L, Gu Q (2002) Dynamic support vector machines for non-stationary time series forecasting. Intell Data Anal 6(1):67–83

A Case Study of Western Maharashtra with Detection and Removal of Cloud Influence: Land Use and Land Cover Change Detection

Renuka Sandeep Gound and Sudeep D. Thepade

Abstract The Western Ghats (Sahyadri Range) and Western Coastal (Konkan) regions are India's most important units of Maharashtra State. This area is mainly comprised Dapoli, Mahabaleshwar, Pune, Kolhapur, Satara, Sangli, Raigad, Ratnagiri, Chiplun, etc. This case study aims to observe the land use and land cover change detection after heavy rainfall of three days duration from July 22nd to July 24th 2021. This Disaster caused many landslides in Western Maharashtra. Images of Land Cover for study are extracted from Sentinel-2 MSI from 2021 to 2022. Minimum distance and Spectral angle mapping classifiers are used to generate the classification map. Classification of Cloud and Non-Cloud area is performed with a Maximum Likelihood classifier. Cloud influence is removed by using Reference Image acquired through Sentinel-2 MSI. The total area utilized for the study is 12056.04 km^2 and six LULC classes are used in the classification. Classification Results obtained exhibit the change in Built-up Area from 2086.72 to 1289.51 km^2, Agriculture sector from 1547.42 to 2115.78 km^2, Plantation from 3999.67 to 4998.79 km^2, Forest Area from 495.82 to 1600.25 km^2, Bare-Soil from 3027.59 to 986.91 km^2, Water Bodies from 898.82 to 1064.8 km^2. Findings and analysis obtained from the statistical analysis show the need for strategies to be designed for continual LULC in Western Maharashtra.

Keywords Land use and land cover (LULC) · A multispectral imagery · SENTINEL-2 (MSI) · Dark object substraction1 (DOS1) · Maximum likelihood classification (MLC) · Normalized difference vegetation index (NDVI)

R. S. Gound (✉)
Sr. Assistant Professor (Computer Sci. & Engineering), New Horizon College of Engineeing, Bangalore, Karnataka, India
e-mail: renuka060182@gmail.com

S. D. Thepade
Professor, Computer Engineering, Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India
e-mail: sudeep.thepade@pccoepune.org

1   Introduction

Land use and Land cover are the two terminologies that can be used interchangeably [1, 2]. Biophysical characteristics of Earth's Surface are referred to as Land Cover. Land cover is typically classified as Waterbodies, Vegetation regions, Baren Soil, Built-ups (developed by humans), etc. This Classification information is needed in many societal applications as well as scientific purposes. Land cover information is often used to analyze environmental change [3, 4]. To assess the changes in agriculture sectors, flood monitoring, heavy rainfall, landslides, risks of climate change, Land cover maps play an important role to develop action plans or strategies and observation of earth's natural resources according to the growing graph of increased population and their requirements with the available ecosystem [5].

In recent decades, as a result of increased population and development in the economy, drastic changes have been observed in LULC in India [6].

Due to growth in the demands of natural resources, most of the changes occurred in the transformation and management of the Land Use/ Land Covers, causing modifications in enormous parts on the surface of Earth [7].

The graph of the population is growing with an increase in eco-socio demands, which makes the burden on the administration of LULC. This leads to random and unintended alterations in Land Use/ Land Cover [8, 9]. Most of the landslides, disasters like a flood occurs as a result of unplanned and lacks of essential management of different land covers like Forest area, Built-up sectors, or development activities by human, Agriculture Land, etc.

Quality of soil is also getting reduced after Land use/ land cover (LULC) changes due to change in climate, and progressed to modification in hydrological cycle with deterioration in the quality of land, ecosystem, agriculture, etc. [10, 11].

To assign the material available on the Earth's Surface to the LULC class, it is necessary to know the details of land cover with modified land due to land use. It is recommended to utilize the worthwhile tools in GIS and Remote sensing [12]. Recognition of Land use/ Land cover change is the more concerned subject in the field of Remote sensing [13].

It has been accepted that from a long time with various spatial scales, changes in LULC can be identified by using the tools related to Geographical Information System (GIS) and Remote Sensing [14, 15].

It is a difficult task to prepare action plans, design policies, administrative decisions related to resource management, with the map developed for Land Cover. Technical institutions are taking efforts to prepare land cover maps of developing units though there is influenced by various challenges [16].

The more challenging fact is to compare or analyze the results generated by land cover maps as the data may be acquired from different satellites with distant sensing dates, statistics used to validate the findings of Land use/land cover change, classification technique applied, the band combination utilized may have particular spatial resolutions, Map is designed for a special interest from the different organization [17].

Land use/Land cover change detection survey is needed to be exercised and revised regularly to assist the activities of designing strategies and solving environmental challenges [18]. Land Cover maps are generally produced by remote sensing tools and are usually known as classification maps. Development of this classification maps accurately needs expertise process through an experienced user, Land cover images with no interference of climatic noise and specific classification technique [19].

A precise method to validate the findings is to observe the transformation of one land cover class to another class and check the percentage of the area changed may support the development of suitable policies to administrate the alteration of the surface of the Land Cover [20].

Remote sensing and GIS tools are the most promising tools to observe the alterations in LULC maps with more accuracy in diverse applications so many researchers have availed them in soil degradation, deforestation, mislaying in the ecosystem of a specific region, etc. [21].

Western Maharashtra experienced heavy rainfall in the month of July-2021 and caused frequent landslides in this region. Due to this, there was an enormous alteration in this field. This article aims to identify approximate land cover changes in the duration of 2021–2022. The changes may be combined effect of disaster as well as climatic conditions.

2   Study Area

The study area is selected from Western Ghats and the western coastal region in Maharashtra. Maharashtra state has a significant and leading role in the agriculture sector in India. The Latitude and Longitude of Maharashtra State, India, are 19.60, and 75.55, respectively.

Maharashtra falls in a tropical region climate. Western Ghats (Sahyadri Mountains) and the Western Coastal region (Konkan) particularly have the highest rainfalls. The average heavy rainfall received in the Konkan is from 2000 to 2500 mm (80–100 in.), and in the Western Ghats adjacent to the Sahyadri ranges are over 5000 mm (200 in.) [22].

Western Ghats and the Western Coastal region has the worst disaster in the three days from July 22nd to July 24th 2021, due to heavy rainfall, there were frequent landslides in these regions. Most of the villages in these regions have suffered a lot from this disaster. Times of India has one report which highlights this disaster with a comment "The increase in encroachments on natural hill slopes for farming or expansion of habitation, deforestation and other anthropogenic reasons have added to the increased frequency of landslides occurrences in past few decades, particularly in Western Maharashtra and Konkan region" [23, 24]. Approximately 28,000 km^2 area in Maharashtra has been affected in the last few decades due to frequent Landslide occurrences.
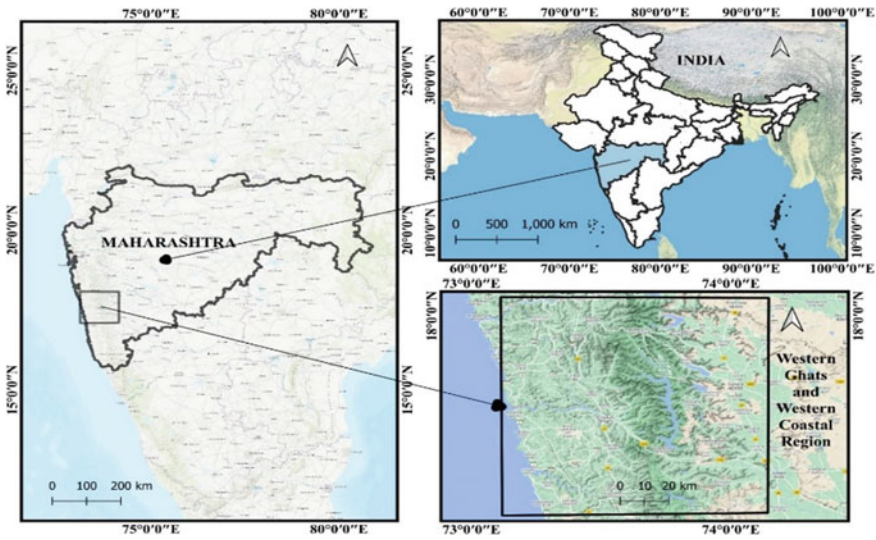
Fig. 1 Location map of Western Ghats and Western Coastal Region, Maharashtra, India

The present article attempts to find an approximate area changed due to the landslides which took place in the month of July-2021. The area change may have the combined effect of meteorological seasonal differences and landslides, from July-2021 to January-2022.

The location of the study area is 73–74° in the East and 18° in the North. Approximately 12,056.04 km^2 area of the Western region is used for the study.

To observe the Land Cover Change Detection, the present study was conducted in the Western Maharashtra Region, India (Fig. 1).

This region especially, in Western Ghats is popularly known for Agriculture Sector as it is the main occupation of around 80% of the population living here. Food Crops mainly included in these divisions are sugarcane, millets, cotton, rice, turmeric, mangoes, banana, etc.

Study Area comprises Satara, Mahabaleshwar, Panchgani, Koyna Nagar, Koyna wild Life Sanctuary (Forest Area), Thseghar (WaterFall), Chandoli National Park, etc. (Western Ghats region) and Chiplun, Dapoli, Guhagar, Ganpatipule, Hedavi, Jaygad, etc. (Western Coastal region).

2.1 Data Product

A Multispectral Imagery, SENTINEL-2 (MSI) is utilized for the acquisition of the required Images needed for study by using https://scihub.copernicus.eu/apihub in Semi-Classification Plugin (SCP) [25] of QGIS. A detailed description of the data product (Study Area) is displayed in Table 1.

Table 1 Details of satellite image

Data product	Spacecraft	Processing level	Sensing date	Physical bands	Resolution (m)
T43QCV	Sentinel-2A	Level-1*C*	07-28-2021 04-04-2021 01-19-2022	*B*2-Blue, *B*3-Red, *B*4-Green	10
				*B*5-Vegetation, *B*7-Vegetation	20
				*B*11-SWIR, *B*12-SWIR	20

3 Materials and Methods

For better results, band images used in the study are pre-processed for atmospheric correction and performed Cloud removal process (Fig. 2) For atmospheric correction, Dark Object Substraction1 (DOS1) is applied.

Three input images (Fig. 2) of the selected portion of Western Maharashtra are considered for identification of the land cover change, Reference Image is acquired on 4-4-2021, Input Image after Heavy Rainfall is acquired on 7-28-2021, to identify Land Change Input Image is acquired on 1-19-2022.

An image of the Landcover acquired on 7-28-2021 is having a cloud cover of 71%. An image with cloud coverage is not used in LULC, so it is needed to generate
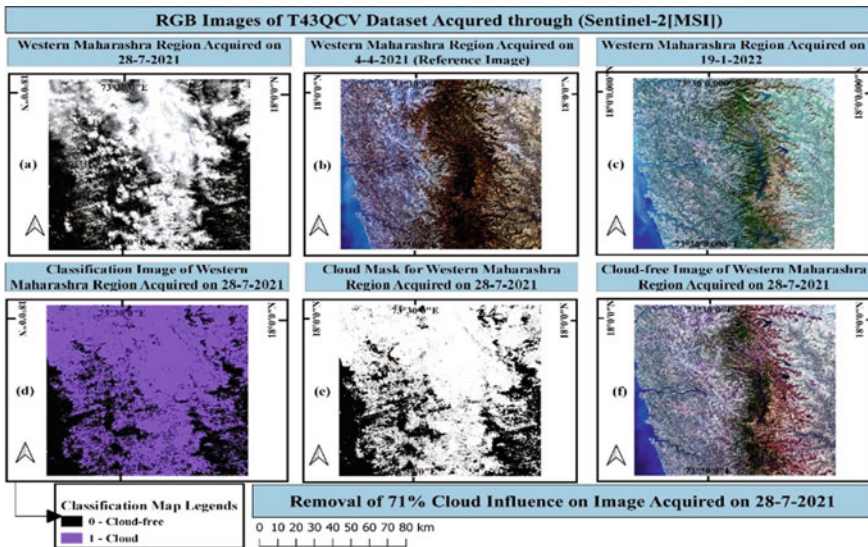
Fig. 2 Input images of selected regions in Western Maharashtra **a** acquired on 7-28-2021 **b** acquired on 4-4-2021(reference image) **c** acquired on 1-19-2022 and removal of cloud influence on the image acquired on 7/28/2021 **d** classification image **e** cloud mask **f** cloud-free image

the cloud-free image for further processing [26]. In preprocessing step classification map of the image is generated with two classes namely cloud area and clear area by applying the Maximum Likelihood Algorithm. Cloud area pixels are removed and replaced with the reference image. Cloud-free image is prepared for further processing.

3.1   Land Use/ Land Cover Detection

Analysis of the material available on Land Cover is carried out by supervised classification. Minimum Distance Classification of SCP is applied over the study area. Minimum Distance Classification [20] is used to assign a pixel to its nearest class. Mean vectors of each class are used for the calculation of the Euclidean distance of each unknown pixel with it. So, the pixel is assigned to the class with the basis on the spectral signature which one is closer to it by using Euclidean Distance measure [23].

The Maximum Likelihood Classification (MLC) Algorithm is commonly used for classification in remote sensing applications. But the time complexity of MLC is very high, it requires more time for computations. Minimum Distance algorithm as compared to MLC performs much better in terms of time required for computations. The land cover used in the study is classified into six classes, Built-up Area, Agriculture Land, Plantation which may include sparse vegetation area as well as trees, shrubs, bushes, grass, etc., Forest area is depicting the area where dense trees are seen, Bare-Soil may contain the rocky area or land which one is not having any vegetation, Water-body represents rivers, lakes, ponds, sea, canals, etc. However, for accurate classification here the NDVI index is calculated and with threshold value obtained, pixels are replaced with the Forest class. This process is applied to Land Covers acquired on 7-28-2021 and 1-19-2022 (Fig. 3).

4   Results and Discussion

Landcover change from July-2021 to January-2022 is determined by comparing pixel by pixel values of classification images with Euclidean Distance difference of Minimum Distance algorithm. The classification report (Table 2) of both the Land Covers used for the study, highlights the Land Cover Class and its area in percentage as well as km^2.

The contribution of area per class depicts that as per classification of the image acquired on 28-7-2021, coverage of the built-up area is 2086.72 km^2 (17.31%), Agriculture land is 1547.42 km^2 (12.84%), Plantation is 3999.67 km^2 (33.18%), Forest Area is 495.82 km^2 (4.11%), Bare-Soil is 3027.59 km^2 (25.11%) and Water-Body is 898.82 km^2 (7.46%) among the total area of 12,056.04 km^2.
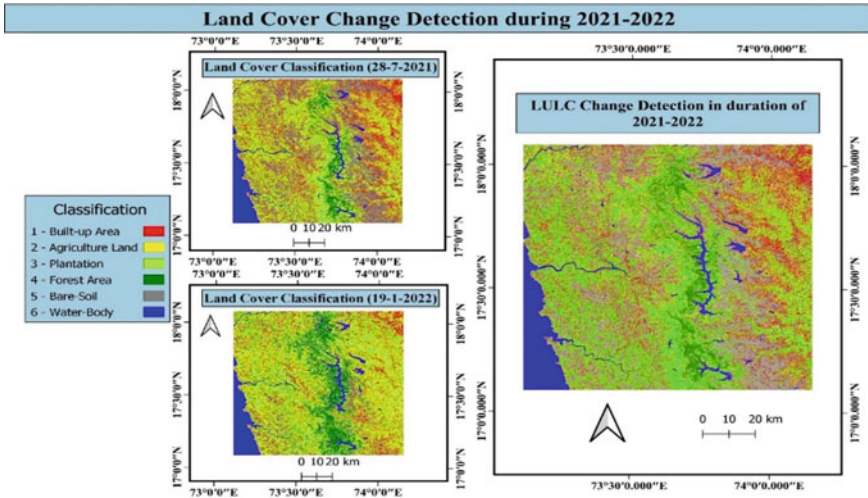
Fig. 3 Land use/land cover change during (2021–2022) based on Sentinel-2 (MSI)

Table 2 Classification report

Land cover classification	C1_7-28-2021		C2_1-19-2022		Change in C1 and C2	
	Percentage (%)	Area (km²)	Percentage (%)	Area (km²)	Percentage (%)	Area (km²)
Built-up area	17.3	2086.72	10.7	1289.51	− 6.6	− 797.21
Agriculture land	12.84	1547.42	17.55	2115.78	4.71	568.36
Plantation	33.18	3999.67	41.46	4998.79	8.28	999.12
Forest area	4.11	495.82	13.27	1600.25	9.16	1104.43
Bare-soil	25.11	3027.59	8.19	986.91	− 16.92	− 2040.68
Water-body	7.46	898.82	8.83	1064.8	1.37	165.98
Total	100	12,056.04	100	12,056.04	0	0

Similarly, we can observe the area percentage of each class of an image acquired on 19-1-2022. Built-up area is 1289.51 km² (10.70%), Agriculture land is 2115.78 km² (17.55%), Plantation is 4998.79 km² (41.46%), Forest Area is 1600.25 km² (13.27%), Bare-Soil is 986.91 km² (8.19%) and Water-Body is 898.82 km² (8.83%).

Classification images prepared (C1_7-28-2021 and C2_1-19-2022) for the study area reveal the alteration in Land Cover during 2021–2022. Computed changes (Fig. 4, Table 2) states that Built-up area is decreased by 6.6%, Agriculture Land is increased by 4.71%, the rise of 8.28% in Plantation, Growth of 9.16% in Forest area, Bare-Soil is reduced by 16.92% and 1.37% inflation in Water-Body area.
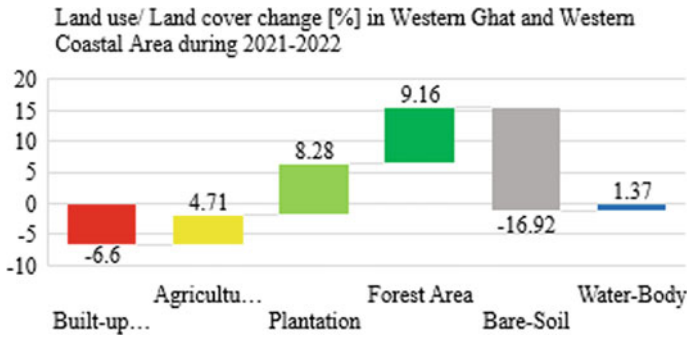
Fig. 4 Land use/land cover change (%)

Due to the frequent landslides that occurred in the duration of 7-22-2021 to 7-24-2021, most of the people lost their homes, damages to various infrastructure facilities have happened, around 1020 villages are affected in the districts of Ratnagiri, Raigad, Satara, Kolhapur and Sangli [22], therefore drastic changes in the different classes can be observed especially in Built-up class.

Land encroachment (Table 3) displays the transformation of areas [km^2] during 2021–2022. Entries shown in Table 3 can be understood as 411.36 km^2 of Built-up area was remains unchanged, but 392.06 km^2 was converted into Agriculture Land, 1107 km^2 is transformed into Plantation, 64.43 km^2 is altered as Forest Area, 86.92 km^2 is converted into Bare-Soil and 24.3 km^2 was seen as waterbody in classification image of the changed Landcover.

For Agriculture Land conversion (124.27 km^2 as Built-up area, 363.28 km^2 as Agriculture Land (No Change), 882.18 km^2 as Plantation, 110.08 km^2 as Forest Area, 48.99 km^2 as Bare-Soil, 18.62 km^2 as Water-Body). Similarly, for other categories also the transformation of land classifications can be interpreted.

Table 3 Land encroachment [km^2] in Western Ghats and Western Coastal area during 2021–2022

Land cover classification	Built-up area	Agriculture land	Plantation	Forest area	Bare-soil	Water-body	Total area
Built-up area	411.36	392.06	1107.65	64.43	86.92	24.3	2086.72
Agriculture land	124.27	363.28	882.18	110.08	48.99	18.62	1547.42
Plantation	40.69	931.55	1938.23	949.48	49.28	90.43	3999.67
Forest area	2.04	22.2	41.35	419.02	5.92	5.28	495.82
Bare-soil	703.91	400.57	1014.18	56.6	789.19	63.14	3027.59
Water-body	7.23	6.12	15.2	0.64	6.6	863.02	898.82
Total	1289.5	2115.78	4998.79	1600.25	986.91	1064.8	12,056.04

Table 4 Land encroachment [%] in Western Ghats and Western Coastal area during 2021–2022

Land cover classification	Built-up area	Agriculture land	Plantation	Forest area	Bare-soil	Water-body
Built-up area	31.90	18.53	22.16	4.03	8.81	2.28
Agriculture land	9.64	17.17	17.65	6.88	4.96	1.75
Plantation	3.16	44.03	38.77	59.33	4.99	8.49
Forest area	0.16	1.05	0.83	26.18	0.60	0.50
Bare-soil	54.59	18.93	20.29	3.54	79.97	5.93
Water-body	0.56	0.29	0.30	0.04	0.67	81.05

Percentage-wise land encroachment (Table 4) determined by classification images can be learned in the same way, consider Plantation (3.16% is transformed into Built-up area, 44.03% as Agriculture Land, 38.77% is as Plantation (no change), 59.33% as Forest Area, 4.99% as Bare-Soil, 8.49% in Water-Body).

5   Conclusion

The study was conducted in the Western Ghats (Sahyadri Range) and Western Coastal (Konkan) regions of Maharashtra State, India to observe the alteration in land use and land cover after heavy rainfall of three days duration from July 22nd to July 24th 2021. Due to the frequent landslides that occurred in this duration, there was an enormous change in Land Cover, Images of Land Cover for study are extracted from Sentinel-2 MSI from 2021 to 2022. Preprocessing of images was carried out to reduce atmospheric scattering by using Dark Object Subtraction 1(DOS1) and to remove the influence of Cloud coverage of 71% by using Maximum Likelihood Classifier (MLC) and reference Image acquired through Sentinel-2 MSI. The total area utilized for the study is 12,056.04 km^2 and six LULC classes are used in the classification. Classification Results obtained exhibit that Built-up Area is reduced by 6.6% (797.21 km^2) which may be the effect of Landslides in this region, Agriculture sector is increased by 4.71% (568.36 km^2), Plantation or sparse vegetation has shown a rise of 8.28% (999.12 km^2), Forest Area is increased by 9.16% (1104.43 km^2) may be caused due to increase in dense vegetation, Bare-Soil is decreased by 16.92% (2040.68 km^2) and increase in Water Bodies by 1.37% (165.98 km^2). The present study has attempted to discover the land change that occurred after the disaster in the month of July-2021 which was more destructive and affected around 1020 villages of western Maharashtra. The findings of the analysis show the approximate land change from July 28th 2021 to January 19th 2022.

References

1. Rawat JS, Kumar M (2015) Monitoring land use/cover change using remote sensing and GIS techniques: a case study of Hawalbagh block, district Almora, Uttarakhand, India. Egypt J Remote Sens Space Sci 18(1):77–84. ISSN: 1110-9823. https://doi.org/10.1016/j.ejrs.2015.02.002.

2. Dimyati M, Mizuno K, Kitamura T (1996) An analysis of land use/cover change using the combination of MSS Landsat and land use map: a case study in Yogyakarta. Indonesia Inter J Rem Sen 17:931–944

3. Ngondo J, Mango J, Liu R, Nobert J, Dubi A, Cheng H (2021) Land-use and land-cover (LULC) change detection and the implications for coastal water resource management in the Wami-Ruvu Basin, Tanzania. Sustainability 13:4092. https://doi.org/10.3390/su13084092

4. Naikoo MW, Rihan M, Ishtiaque M, Shahfahad (2020) Analyses of land use land cover (LULC) change and built-up expansion in the suburb of a metropolitan city: spatio-temporal analysis of Delhi NCR using landsat datasets. J Urban Manage 9(3):347–359. ISSN: 2226-5856. https://doi.org/10.1016/j.jum.2020.05.004

5. Vivekananda GN, Swathi R, Sujith AVLN (2021) Multi-temporal image analysis for LULC classification and change detection. Eur J Remote Sens 54(sup2):189–199. https://doi.org/10.1080/22797254.2020.1771215

6. Moulds S, Buytaert W, Mijic A (2018) A spatio-temporal land use and land cover reconstruction for India from 1960–2010. Sci Data 5:180159. https://doi.org/10.1038/sdata.2018.159

7. Rahaman S, Kumar P, Chen R, Meadows ME, Singh RB (2020) Remote sensing assessment of the impact of land use and land cover change on the environment of Barddhaman District, West Bengal, India. Front Environ Sci 8. https://www.frontiersin.org/article/10.3389/fenvs.2020.00127

8. Reis S (2008) Analyzing land use/land cover changes using remote sensing and GIS in Rize, North-East Turkey. Sensors 8(10):6188–6202

9. Seto KC, Woodcock CE, Song C, Huang X, Lu J, Kaufmann RK (2002) Monitoring and use change in the Pearl River Delta using Landsat TM. Int J Remote Sens 23(10):1985–2004

10. Roy A, Inamdar AB (2019) Multitemporal land use land cover (LULC) change analysis of a dry semi-arid river basin in western India following a robust multisensor satellite image calibration strategy. Heliyon 5:e01478. https://doi.org/10.1016/j.heliyon.2019.e01478

11. Bajocco S, De Angelis A, Perini L, Ferrara A, Salvati L (2012) The impact of land use/land cover changes on land degradation dynamics: a Mediterranean case study. Environ Manage 49:980e989

12. Olokeogun OS, Iyiola K, Iyiola OF (2014) Application of remote sensing and GIS in land use/land cover mapping and change detection in Shasha forest reserve, Nigeria. In: ISPRS—international archives of the photogrammetry, remote sensing and spatial information sciences, vol. XL8, pp 613–616. https://doi.org/10.5194/isprsarchives-XL-8-613-2014

13. Mukherjee S. Shashtri S. Singh CK. Srivastava PK, Gupta M (2009) Effect of canal on land use/land cover using remote sensing and GIS. J Indian Soc Remote Sens

14. Twisa S, Buchroithner MF (2019) Land-use and land-cover (LULC) change detection in Wami River Basin, Tanzania. Land 8:136. https://doi.org/10.3390/land8090136

15. Dewan A, Corner R (2014) Dhaka megacity: geospatial perspectives on urbanisation, environment, and health. Springer Science and Business Media, Berlin

16. Saah D, Tenneson K, Matin M, Uddin K, Cutter P, Poortinga A, Nguyen QH, Patterson M, Johnson G, Markert K, Flores A, Anderson E, Weigel A, Ellenberg WL, Bhargava R, Aekakkararungroj A, Bhandari B, Khanal N, Housman IW, Potapov P, Tyukavina A, Maus P, Ganz D, Clinton N, Chishtie F (2019) Land cover mapping in data scarce environments: challenges and opportunities. Front Environ Sci 7:150. https://doi.org/10.3389/fenvs.2019.00150

17. Spruce J, Bolten J, Mohammed IN, Srinivasan R, Lakshmi V (2020) Mapping land use land cover change in the lower Mekong Basin from 1997 to 2010. Front Environ Sci 8. https://www.frontiersin.org/article/10.3389/fenvs.2020.00021

18. Al-Saady Y, Merkel B, Al-Tawash B, Al-Suhail Q (2015) Land use and land cover (LULC) mapping and change detection in the Little Zab River Basin (LZRB), Kurdistan Region, NE Iraq and NW Iran. FOG Freiberg Online Geosci 43:1–32
19. Rwanga S, Ndambuki J (2017) Accuracy assessment of land use/land cover classification using remote sensing and GIS. Int J Geosci 8:611–622. https://doi.org/10.4236/ijg.2017.84033
20. Kafi KM et al. (2014) An analysis of LULC change detection using remotely sensed data; a case study of Bauchi City. IOP Conf Ser: Earth Environ Sci 20:012056
21. Joshi S, Rai N, Sharma R, Baral N (2021) Land use/land cover (LULC) change in suburb of Central Himalayas: a study from Chandragiri, Kathmandu. J For Environ Sci 37(1):44–51. https://doi.org/10.7747/JFES.2021.37.1.44
22. Wikipedia contributors (2022) 2021 Maharashtra floods. Wikipedia. Retrieved 28 Feb 2022, from https://en.wikipedia.org/wiki/2021_Maharashtra_floods
23. Gound RS, Thepade SD (2021) Removal Of cloud and shadow influence from remotely sensed images through LANDSAT8/OLI/TIRS using minimum distance supervised classification. Indian J Comput Sci Eng 12(6):1734–1748
24. Dighe S (2021) Encroachment on slopes, deforestation to blame for landslides in Maharashtra: experts. The Times of India. Retrieved 28 Feb 2022, from https://timesofindia.indiatimes.com/city/pune/encroachment-on-slopes-deforestation-to-blame-for-landslides-in-maharashtra-experts/articleshow/84936211.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst
25. Congedo L (2021) Semi-automatic classification plugin: a python tool for the download and processing of remote sensing images in QGIS. J Open Source Softw 6(64):3172. https://doi.org/10.21105/joss.03172
26. Gound RS, Thepade SD (2021) Removing haze influence from remote sensing images captured with airborne visible/infrared imaging spectrometer by cascaded fusion of DCP, GF, LCC with AHE. In: 2021 International conference on computing, communication, and intelligent systems (ICCCIS), pp 658–664. https://doi.org/10.1109/ICCCIS51004.2021.9397060

A Survey on Deep Learning Enabled Intrusion Detection System for Internet of Things

Huma Gupta, Sanjeev Sharma, and Sanjay Agrawal

Abstract Internet of Things (IoT) has led the world to grow at a revolutionary rate that is bringing profound impact in various domains without human intervention. Although IoT networks are enhancing different application domains, they are vulnerable and possess different cyber threats to its adoption and deployment. Thus, the continuously growing threats demand security measures that can lead IoT adoption without risk around the globe. Also, the security methods that are deployed for traditional networks lack security provisioning for IoT networks as these are resource constraint, deploy usage of diverse protocols, therefore existing security schemes demand to be evolved to be adopted for the security of IoT systems. As the IoT ecosystem grows rapidly with time and devices are connected via wireless sensor-based connections, deploying intelligent attack detection methods are one of the promising techniques which can detect different cyber-attacks and monitor IoT networks intelligently. Thus, the intrusion detection systems (IDS) play an important role in the security of IoT networks and evolving IDS with the application of deep learning networks leads to the generation of more powerful security systems that can be capable of detecting even zero birth attacks and making IoT networks more adaptable. The aim of this paper is to provide a comprehensive survey of IDS in IoT networks which has evolved in past years with the usage of machine learning and resulted in improving the performance of IDS. The main objective of this survey is to thoroughly review different deep learning methods for IDS.

Keywords Internet of things security · Attacks · Intrusion detection system · Deep learning

H. Gupta (✉) · S. Sharma
Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India
e-mail: humajain@gmail.com

S. Sharma
e-mail: sanjeev.rgtu@gmail.com

S. Agrawal
National Institute of Technical Teachers' Training and Research, Bhopal, India
e-mail: sagrawal@nitttrbpl.ac.in

1 Introduction

The Internet of Things is the most rapidly growing computing environment that connects different computing devices via the Internet and improves day-to-day life with ubiquitous computing environments. This data-driven communication among different computing devices such as sensors and actuators is evolving very fast and leading in the world as a pervasive computing model [1]. As per the Gartner review, IT Services for the IoT market will represent a 58 billion dollar opportunity in 2025 [2]. This is the most evolving paradigm that is giving rise to smart computing environments such as smart cities, smart homes, and smart service industries for different applications. The entire essence of this computing environment lies in communication via radio frequency identification (RFID) tags which, through unique addressing schemes, are able to interact with each other and allow communication among heterogeneous devices over a variety of protocol standards. However, this communication over the Internet is susceptible to various security attacks that range over wide categories such as denial of service, distributed denial of services, routing attacks, etc., also mentioned in [3]. These attacks are potential threats to IoT networks and may result in the loss of IoT services.

The IoT devices are evolving and becoming smarter with wide usage in diverse domains such as domestic, health, education, business, and agriculture. The confidentiality, integrity, and availability for IoT networks is required to be preserved against different attacks and enhanced security measures must be adopted to secure the IoT ecosystem [3]. The vulnerabilities of IoT systems cannot be neglected and demands measures to protect them against different attacks. Therefore, different existing solutions for securing conventional networks demand enhancement to integrate those techniques for IoT networks. Consequently, more attention is expected from researchers to keep monitoring the attack vectors and analyzing these attacks for their timely detection and prevention.

This paper focuses on providing a study on intrusion detection systems for IoT networks. The rest of the paper is organized as follows: In Sect. 2 various security requirements and attacks in an IoT network are presented. Section 3 details intrusion detection systems and their classifications. In Sect. 4, different intrusion detection systems for IoT using deep learning are briefed and compared. In Sect. 5, conclusions and future discussions are presented.

2 IoT Security Requirements and Attacks

IoT networks communicate in heterogeneous settings, utilizing wireless protocols, thus building attack proof IoT systems is the biggest challenge in deploying device based computing in various application domains. Being web-enabled smart devices, the nodes in the IoT ecosystem continuously transfer huge data using hardware

communicating devices without human intervention. Thus these networks are susceptible to various web-based threats as well as new threats that can sense the weakness of such environments such as distributed denial of service attacks. The most censorious example is the Mirari attack which is a massive malware attack that infects smart devices and turns them into an army of "zombies" [4]. As a result, the IoT ecosystem is demands more security than traditional systems, and it requires a stronger assault defense mechanism. The constraints such as memory capacity, battery life, computational power, and network bandwidth limits make building security procedures for IoT systems more difficult than designing security mechanisms for traditional networks. Undefined technical standards and protocols for the deployment of IoT systems are another difficulty for IoT systems, implying that the IoT security mechanism must adhere to a distinct set of protocols. Also the wide attack vector imposes threat to different layers of IoT architecture [5]. As IoT architecture is composed of three-layer architecture depicted below, each layer deploys a heterogeneous set of protocols that demands security solutions that must consider these different protocols as well as being capable of detecting attacks to these layers.

2.1 Classification of Attacks

The attacks in IoT environments are continuously evolving as these networks are scaling up. These attacks can be classified into passive attacks and active attacks [6]. The passive attack acts silently and steals private information without altering functioning in the devices and, therefore detection of such attacks is challenging. Whereas, the active attacks alter the operations in devices that can be easily traced in system logs. Active attacks can also be further categorized into jamming, flooding, Denial-of-Service (DoS), wormhole, sinkhole, blackhole, and Sybil types.

The IoT architecture is composed of three sub layers: perception layer, network layer, and application layer. Each layer has its protocol standard for operations and each layers is imposed to various attacks. The physical hardware layer is the perception layer. It's in charge of keeping sensors and actuators in communication with one other. RFID, 6LowPAN, and Bluetooth are among the communication standards used. The network layer is in charge of routing data between different devices using protocols GSM, IPv6, and others. The application layer, also known as the software layer, is the topmost layer, and it presents a user interface with business logic implementation. The authors in [6, 7] provide details of different attack vectors as per the architecture of IoT systems. Some common attacks are distributed denial of service, jamming, Trojan, interference, an Sybil attack [8].

2.2 Requirements of IDSs in IoT Networks

With the tremendous growth of IoT networks, the need for security is getting enhanced day by day as the model is getting complex in various domains. The growing infrastructure is raising vulnerabilities in the adoption due to the broadcasting of data communication among the IoT devices making them an easier target for intrusions. Though the communication in the IoT can be secured using different cryptographic techniques for confidentiality, authentication, and authorization, these techniques do not ensure attack-free deployment of IoT. As the number of IoT devices is exponentially increasing, the number of threats and vulnerabilities also rises, which demands a second line of defense to protect these devices. To achieve this extended line of defense Intrusion detection systems are the most preferred solutions. The IDS monitors the network activities within a heterogeneous network and detects malicious activities by generating alarms. The concept of IDS for detecting intrusions and securing network systems has evolved tremendously in the past 25 years [9]. During these years, researchers have proposed different methods that could detect different malicious attacks by performing analysis on captured network traffic. In 1987, Denning [10] proposed an intrusion detection model that could detect malicious behavior for a system by comparing deviation from normal behavior. By then, there are many researchers who have worked in different networks to detect attacks in the systems. These different contributions in the field of IDS are discussed in the next sections of this paper.

An IDS deployed for the conventional system are not fully suitable for IoT systems as the IoT systems demand computation under the rigid condition of light processing devices, high volume data, fast real-time response, and different protocols such as IPv6 over-Long Range Wide Area Network (6LoWPAN) protocol, CoAP, and Message Queuing Telemetry Transport (MQTT) are not a protocol to be adopted in traditional communication networks. Thus, the development of IDS that fulfills the above design criteria is required for the IoT environment. IoT security and mitigation approaches are continuously evolving issues that must be focused deeply and efforts must be made to provide a detection mechanism that can operate efficiently in stringent conditions and provide sections of attacks with high accuracy, low false positives rate, minimum processing time, and processing overhead.

3 IDS for IoT Systems

An IDS is a tool that monitors networks, analyzes user information and traffic to discover the security vulnerabilities and detect intrusion within it. The operations in any detection process involve the monitoring stage where the entire network is monitored periodically via different sensors, the Analysis stage where the collected data is analyzed using different pattern matching and identification techniques to assess the presence of vulnerabilities. Finally, the detection stage deploys different

Fig. 1 IDS process
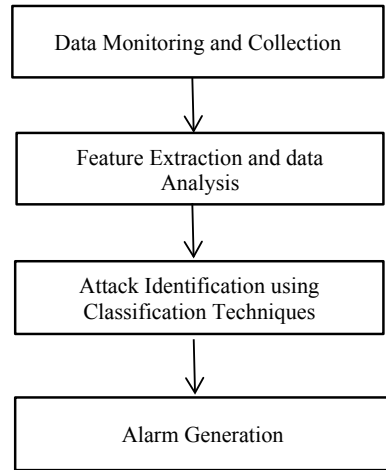
detection techniques to detect intrusion and generate an alarm. This entire process of intrusion detection is depicted in Fig. 1.

Many researchers have given surveys on the evolving nature of IDS with its classification in [11]. The implementation of IDS can be classified on the basis of its deployment environment and selection of methodologies that can be used to detect the intrusion and attacks. On the basis of deployment, IDS are classified as host based IDS (HIDS) and network-based IDS NIDS. A HIDS is designed to deploy in a single node in the network and protect it from attacks that can raise harm to its applications and data [12, 13]. Network intrusion detection systems are deployed within the network where it analyzes the flow of traffic in the network to detect intrusion [14]. A NIDS is deployed intelligently within networks that passively inspect traffic traversing the devices on which devised. A NIDS can be hardware or software-based systems depending upon components types who perform analysis, as it can be either software agents or hardware devices. Snort [15] is one of the most commonly used software-based IDS. Also, as the network grows in size, the traffic volumes also get massive, in such cases, cyber-attacks also get smarter and sophisticated, thus challenges in IDS design grows.

IDS can be further classified into signature-based IDS, anomaly-based IDS, and specification-based IDS. Signature-based IDS uses information of known attacks and patterns of malicious activities to detect intrusions. During the detection process, the pattern in the traffic is matched with stored patterns in the database, if matched alarms are generated. Such IDS demands a high cost of maintaining a database of known attacks, thus previous knowledge of attacks is mandatory. In anomaly-based intrusion detection technique traffic is analyzed to detect abnormal behavior that leads to generating the alarm.

A divergence from regular data patterns formed using data from normal users and activities can be used to detect abnormal activity. Furthermore, the IDS may be a hybrid approach that can ensemble two or more types of IDS with an intention to drive

maximum advantage of an individual approach. A hybrid model is one that combines anomaly and signature-based IDSs to provide a better compromise between storage and computing costs while reducing false positive alarms. Due to its successful detection and easier operation, Hybrid IDS is commonly used in the majority of systems nowadays, as detailed by author in [16]. Furthermore, in recent years researchers have adopted many computing techniques that can be used in designing that touches different aspect of design such as the computational complexity, execution time, and detection time. Such modeling of IDS techniques such as statistical, knowledge-based, data mining-based, machine learning-based, and so on, evolved with different computing trends. Currently, these computing models can also be ensemble with the goal of designing hybrid models that may reach better IDS performance [17].

4 Deep Learning for Intrusion Detection System

IDS for IoT networks are in demand due to growing IoT adoption and increasing attack sets. In IoT environments that are distributed, constrained in resources, and have limited processing capability than conventional networks, demands IDS to be light weighted and meet the requirement of 6LoWPAN, CoAP, RPL, etc., networks. Over the past years researchers have evolved contributions for IDS from some early approaches such as mining, knowledge based, expert systems based to machine learning-based IDS. The authors in [18–20], have surveyed different popular trends for the analysis and attack classification, in the design of IDS. In recent years machine learning and deep learning are the most desirable techniques that help users classify attacks using feature extraction and analysis. The authors in [21] present the survey of machine learning and deep learning approaches in trends. The common machine learning that is being adopted so far for IDS in IoT is support vector machine, decision tree, random forest, k-mean clustering, and ensemble learning. Over the years these machine learning approaches are being adopted to classify attacks and intrusions but with the increasing data and complexity of computation, the efficiency of machine learning faces challenges in achieving desired performance.

Over the machine learning algorithm, the deep learning algorithm provides superior performance over the large dataset with high dimensionality. Deep learning methods introduce multiple nonlinear layers of computations for feature extraction and with pattern-based analysis. The deep learning techniques are inspired by human brain with multiple neurons for making decisions collectively. It is a subfield of machine learning that can be classified into supervised learning, unsupervised learning and the combination of these learning types, which is called hybrid learning. Table 1 lists various deep learning techniques that can be deployed for IoT security models [22]. In the rest of the section we have highlighted the deep learning based intrusion detection system for IoT systems.

Recent advancement in deep learning enabled researchers to present intrusion detection schemes based on deep learning techniques mentioned in the previous section. The authors in [23, 24] have proposed CNN-based detection schemes for

Table 1 Deep learning techniques

Deep learning technique	Example
Supervised learning (discriminative),	Convolutional neural networks (CNN), Recurrent neural networks (RNNs)
Unsupervised learning (generative learning)	Autoencoders (AEs), Restricted Boltzmann machines (RBMs), Deep belief networks (DBNs)
Hybrid learning	Generative adversarial networks (GAN), Ensemble of DL networks (EDLNs)

feature extraction which results in significant performance of the classifier. In addition to CNN, [24] also adopts transient search optimization for complex problems. Further the model is evaluated using three IoT IDS datasets, KDDCup-99, NSL-KDD, and BoT-IoT where the applied model is verified with the improvement in classification accuracy.

Thamilarasu et al. [24] uses a deep belief network (DBN) to fabricate the feedforward deep neural network (DNN) as the perceptual learning model. The proposed learning model is evaluated for the detection of various attacks such as sinkhole attack, distributed denial-of-service (DDoS) attack, blackhole attack, opportunistic service attack, and wormhole using network simulation and testbed implementation. Khan et al. [25] have proposed a DNN-based intrusion detection system for MQTT-enabled IoT smart systems. This work detects different attacks which are attracted by MQTT-based communication. The proposed IDS is compared with other traditional machine learning algorithm such as Naive Bayes (NB), Random Forest (RF), *k*-Nearest Neighbors (KNN), Decision Tree (DT), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRUs). Muder et al. [26] have proposed an artificial intelligent intrusion detection system for fog based IoT. The proposed techniques consist of two engines, one for traffic processing engine and another is classification engine. This design leads to intelligent multilayer IDS that enhance the stability of the neural network response. As the two detection layers are capable of detecting hidden attacks too. Nagaraj et al. [27] designed a deep belief network (DBN)-based IDS. The DBN is stacked with many layers, including restricted boltzmann machine (RBM) which are designed with many stages that allows classification of breaches in IoT systems.

Zhong et al. [28] have proposed a sequential learning based intrusion detection system which deploys gated recurrent unit (GRU) and Text-CNN deep learning methods. Shahriar et al. [29] proposed a deep learning-based IDS using generative adversarial networks (GAN) which is a promising hybrid deep learning technique. The authors have compared the performance of S-IDS and a GAN supported IDS, which performs better than standalone IDS. Qaddoura et al. [30] designed a multi-layer classification approach which handles the problem of imbalanced dataset and detects intrusion with its type. It uses SMOTE oversampling technique for oversampling of minority dataset that generated a balanced dataset which is further used to train single-hidden layer feed-forward neural network (SLFN) and a two layer LSTM. Lansky et al. [31] presented a systematic review of the deep learning-based

Table 2 Summary of deep learning based IDS

References	Methodology	Dataset/stimulation	Accuracy/precision
Al-Garadi et al. [20]	CNN	Malimg dataset	99%
Fatani et al. [23]	CNN	KDDCup-99, NSL-KDD, and BoT-IoT	99.994%
Thamilarasu et al. [24]	DNN	Stimulation	99.5%
Khan et al. [25]	DNN	MQTT-IoT-IDS2020	99.9333% in uniflow, 99.788% in biflow, 94.996% in packet flow
Muder et al. [26]	RNN	NSL KDD	92.18%
Nagaraj et al. [27]	DBN	Stimulation	99%
Zhong et al. [28]	Text-CNN and GRU	KDD-99 and ADFA-LD	95% $F1$ score
Shahriar et al. [29]	GAN	NSL KDD-99	98%
Qaddoura et al. [30]	SLFN + LSTM	IoTID20	98.5%

IDS schemes according to the type of deep learning network utilized in them. The authors in the surveys have reviewed more than hundred papers of different deep learning techniques such as auto encoders, RBM, RNN, DBN, DNN, and CNN. The papers also covered the comparison of these IDS under different evaluation parameters. Later the paper covers different challenges, issues with deep learning-based IDS development. This comprehensive survey has provided a detailed view of deep learning implementation using a massive dataset and training process which is further in process of improvement with training speed.

A short comparison on above discussed deep learning based IDS is summarized in Table 2 with focus on methodology adopted and the performance achieved.

5 Conclusion, Challenges, and Future Discussion

The survey focuses on study of deep learning techniques for intrusion detection system for IoT networks. With the evolving nature of attacks on IoT networks, the deployment of IDS is introduced with the wide classification of IDS and types. Further, we have discussed different IDS which are being proposed by researchers for IoT networks. As traditional learning approaches need to be optimized with the growing nature of this network, we have discussed various IDS which have benefited from deep learning techniques. Also, these IDS are compared under different evaluation metrics and datasets used. This survey investigates the importance of deep learning schemes and proofs to be a promising area for researchers. Based on different IDS surveyed using different deep learning approaches, the adoption of the dataset in the training of learning models is commonly performed using conventional

network datasets such as KDD99, NSL-KDD which contain old traffic data which is hard to be mapped for the current IoT environment. Thus the approaches of IoT IDS must be trained and evaluated using IoT datasets that are capable of mapping IoT traffic and attack identification. Also, devices in IoT are resource-constrained devices. The resources, such as memory, computation, and energy, are limited and impose challenges in the adoption of learning based detection models for real-time onboard implementation. Thus developing learning based frameworks that can efficiently reduce computational complexity must be the significant goal for researchers. Further, the application of ensemble deep learning classifiers is a promising area that can lead to building IDS with higher accuracy and precision.

References

1. Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. Comput Netw 54(15):2787–2805. https://doi.org/10.1016/j.comnet.2010.05.010
2. National Intelligence Council (2008) Disruptive civil technologies—six technologies with potential impacts on US Interests out to 2025. Conference Report CR 2008-07. Retrieved from http://www.dni.gov/nic/NIC_home.html
3. Grammatikis PIR, Sarigiannidis PG, Moscholio ID (2019) Securing the internet of things: challenges, threats and solutions. Internet Things 5:41–70. https://doi.org/10.1016/j.iot.2018.11.003
4. Kolias C, Kambourakis G, Stavrou A, Voas J (2017) DDoS in the IoT: mirai and other botnets. Computer 50(7):80–84. https://doi.org/10.1109/MC.2017.201
5. Österberg BP, Song H (2020) Security of the internet of things: vulnerabilities, attacks, and countermeasures. IEEE Commun Surv Tutorials 22(1):616–644. https://doi.org/10.1109/COMST.2019.2953364
6. Mosenia A, Jha NK (2017) A comprehensive study of security of internet-of-things. IEEE Trans Emerg Top Comput 5(4):586–602. https://doi.org/10.1109/TETC.2016.2606384
7. Gupta H, Sharma S (2021) Security challenges in adopting internet of things for smart network. In: 10th IEEE international conference on communication systems and network technologies (CSNT), pp 761–765. https://doi.org/10.1109/CSNT51715.2021.9509698
8. Zhang K, Liang X, Lu R, Shen X (2014) Sybil attacks and their defenses in the internet of things. IEEE Internet Things J 1(5):372–383. https://doi.org/10.1109/JIOT.2014.2344013
9. Butun I, Morgera SD, Sankar R (2014) A survey of intrusion detection systems in wireless sensor networks. IEEE Commun Surv Tutorials 16(1):266–282. https://doi.org/10.1109/SURV.2013.050113.00191
10. D. E. Denning (1987). An Intrusion Detection Model. In: IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pp. 222–232. doi: https://doi.org/10.1109/TSE.1987.232894.
11. Liao HJ, Lin CH, Lin YC, Tung KY (2013) Intrusion detection system: a comprehensive review. J Netw Comput Appl 36(1):16–24. https://doi.org/10.1016/j.jnca.2012.09.004
12. Creech G, Hu J (2014) A semantic approach to host-based intrusion detection systems using contiguous and discontinuous system call patterns. IEEE Trans Comput 63(4):807–819. https://doi.org/10.1109/TC.2013.13
13. Gautam SK, Om H (2016) Computational neural network regression model for host based intrusion detection system. Perspect Sci 8:93–95. https://doi.org/10.1016/j.pisc.2016.04.005
14. Maciá-Pérez F, Mora-Gimeno FJ, Marcos-Jorquera D, Gil-Martínez-Abarca JA, Ramos-Morillo H, Lorenzo-Fonseca I (2011) Network intrusion detection system embedded on a smart sensor. IEEE Trans Ind Electron 58(3):722–732. https://doi.org/10.1109/TIE.2010.2052533

15. Team S (2017) Snort-network intrusion detection and prevention system. Retrieved from https://www.snort.org/
16. Smys S, Basar A, Wang H (2020) Hybrid intrusion detection system for internet of things. J ISMAC 2:190–199
17. Aburomman AA, Reaz MBI (2017) A survey of intrusion detection systems based on ensemble and hybrid classifiers. Comput Secur 65:135–15. https://doi.org/10.1016/j.cose.2016.11.004
18. Ioulianou P, Vasilakis V, Moscholios I, Logothetis M (2018) A signature-based intrusion detection system for the internet of things .In: Information and communication technology forum
19. Chaabouni N, Zemmari MA, Sauvignac C, Faruki P (2019) Network intrusion detection for IoT security based on learning techniques. IEEE Commun Surv Tutorials 21(3):2671–2701. https://doi.org/10.1109/COMST.2019.2896380
20. Al-Garadi MA, Mohamed A, Al-Ali AK, Du X, Ali I, Guizani M (2020) A survey of machine and deep learning methods for internet of things (IoT) security . IEEE Commun Surv Tutorials 22(3):1646–1685. https://doi.org/10.1109/COMST.2020.2988293
21. Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE (2017) A survey of deep neural network architectures and their applications. Neurocomputing 234:11–26. https://doi.org/10.1016/j.neucom.2016.12.038
22. Anand A, Rani S, Anand D, Aljahdali HM, Kerr D (2021) An efficient CNN-based deep learning model to detect malware attacks (CNN-DMA) in 5G-IoT healthcare applications. Sensors 21:6346. https://doi.org/10.3390/s21196346
23. Fatani A, Abd Elaziz M, Dahou A, Al-Qaness MAA, Lu S (2021) IoT intrusion detection system using deep learning and enhanced transient search optimization. IEEE Access 9:123448–123464. https://doi.org/10.1109/ACCESS.2021.3109081
24. Thamilarasu G, Chawla S (2019) Towards deep learning driven intrusion detection for the internet of things. Sensors. https://doi.org/10.3390/s19091977
25. Khan MA, Jan SU, Ahmad J, Jamal SS, Shah AA, Pitropakis N, Buchanan WJ (2021) A deep learning-based intrusion detection system for MQTT enabled IoT. Sensors. https://doi.org/10.3390/s21217016
26. Almiani M, AbuGhazleh A, Al-Rahayfeh A, Atiewi S, Razaque A (2020) Deep recurrent neural network for IoT intrusion detection system. Simul Model Pract Theory 101:102031. https://doi.org/10.1016/j.simpat.2019.102031
27. Balakrishnan N, Rajendran A, Pelusi D, Ponnusamy V (2021) Deep belief network enhanced intrusion detection system to prevent security breach in the internet of things. Internet Things 14. https://doi.org/10.1016/j.iot.2019.100112
28. Zhong M, Zhou Y, Chen G (2021) Sequential model based intrusion detection system for IoT servers using deep learning methods. Sensors 21(4). https://doi.org/10.3390/s21041113
29. Shahriar MH, Haque NI, Rahman MA, Alonso M (2020) G-ids: generative adversarial networks assisted intrusion detection system. In: IEEE 44th annual computers, software, and applications conference (COMPSAC), pp 376–385. https://doi.org/10.3390/electronics11040524
30. Qaddoura R, Al-Zoubi M, Faris H, Almomani I (2021) A multi-layer classification approach for intrusion detection in IoT networks based on deep learning. Sensors 21(9):2987
31. Lansky J, Ali S, Mohammadi M, Majeed M, Karim S, Rashidi S, Hosseinzadeh M, Rahmani A (2021) Deep learning-based intrusion detection systems: a systematic review. IEEE Access 9:101574–101599. https://doi.org/10.1109/ACCESS.2021.3097247

Design and Development
of Multithreaded Web Crawler
for Efficient Extraction of Research Data

Poornima G. Naik and Kavita S. Oza

Abstract The semantic Web renders the search, machine friendly rather than human friendly. Semantic Web comprises of a set of principles, various tools, and techniques which enables people to share their content on the Web and participate in a collaborative work for quickly sharing the valuable information on Web. The current research focuses on the combined qualitative and quantitative research methodology to find out the specific requirements of data with respect to research area. The authors have designed and implemented semantic framework for querying the research data in machine readable format employing RDF, FOAF, and ontology semantic Web technologies. The current paper presents a multithreaded Web crawler module for extracting the attributes of interest pertaining to different attributes of research paper such as indexing, journal impact factor, processing charges, volume no, issue no, e-ISSN, and p-ISSN. The model is generic and can be applied to any domain for retrieving the attributes of interest. The research objective is to minimize the time involved in literally going through each journal listed by the search engine and manually finding out the requisite information.

Keywords AJAX · RDF generator · RDF validator · Sematic Web · Web crawler · Web ontology

1 Introduction

Owning to the digitization of data in virtually all sectors, a sea change is experienced in information generation, distribution, and access. In recent years, finding information from the Web is becoming more and more tedious. Further, the data extracted is not directly consumable by human. Searching on the Internet can be compared to

P. G. Naik
CSIBER, Kolhapur, Maharashtra, India
e-mail: pgnaik@siberindia.edu.in

K. S. Oza (✉)
Shivaji University, Kolhapur, Maharashtra, India
e-mail: kso_csd@unishivaji.ac.in

dragging a net across the surface of the ocean. In this scenario, the traditional search engines are becoming obsolete for providing satisfactory solutions for addressing these issues. Research is in progress to build a Web, which is semantically richer than the current one. There is a need for building a powerful artificial intelligent (AI) system incorporating the ability to translate knowledge from different languages. The semantic Web relies heavily on the formal ontologies that structure underlying data for the purpose of comprehensive and transportable machine understanding. Semantic Web technology relies on ontology as a tool for modeling an abstract view of the real world and contextual semantic analysis of documents. Therefore, the success of the semantic Web depends predominantly on the proliferation of ontologies, which requires fast and easy engineering of ontology and avoidance of a knowledge acquisition bottleneck. The major drawback of the current Web is that Web data lacks machine-understandable semantics, so it is generally not possible to automatically extract concepts or relationships from this data or to relate items from different sources. The Web community is attempting to address this limitation by designing a semantic Web. The semantic Web aims to provide data in a format that embeds semantic information and then seeks to develop sophisticated query tools to interpret and combine this information. The result should be a much more powerful knowledge sharing environment than today's Web: instead of posing queries that match text within documents, a user could ask questions that can only be answered via inference or aggregation; data could be automatically translated into the same terminology.

1.1   State of Current Web

The current Web is unstructured which poses a large limitation on aggregating data from different similar resources. The search query results are human readable but are not machine readable. To address this issue, semantic Web aims at rendering structured layer on top of unstructured data repository which facilitates data retrieval through inference mechanism. Much of the research focus on the semantic Web is based on treating the Web as a knowledge base defining meanings and relationships.

1.2   Introduction to Web Crawler

To perform search engine optimization, Web crawlers play a pivotal role. A Web crawler (also known as a Web spider or search engine robot) is a programmed script that browses the World Wide Web in a methodical, automatic manner. This process is called as Web crawling or spidering. Search engines and many Websites make use of crawling as a means of providing the latest data. Web crawlers are generally used to create a copy of all the visited pages for subsequent processing by the search engine that will index the downloaded pages to facilitate fast searches. To put it simply, the

Web crawler or Web spider is a type of bot, or a software program that visits Web sites and reads their pages and other information to create entries for a search engine index.

When the Web spider returns home, the data is indexed by the search engine. All the main search engines, such as Google and Yahoo, use spiders to build and revise their indexes. Moreover, the spiders or crawlers are used for automating maintenance tasks on the Website. For example, the crawlers are used to validate HTML code, gather certain types of data from the Websites, and check links on the Web page to other sites.

The authors have developed a sematic framework for querying the research data. The current research focuses on design and development of multithreaded Web crawler for efficient extraction of research data from the Web which comprises of attributes of interest. A multithread application is implemented in Java for parsing each html document by spawning multiple browser instances for invoking html DOM parser implemented in PHP.

2 Current State of Research in Semantic Web

Semantic Web along with artificial intelligence (AI) is expanding to devices and things in everyday life. This new combination of semantic Web and AI calls for the new approach and a different look on knowledge representation and processing at scale for the semantic Web. A group of researchers from academia and industry discussed fundamental concepts with regards to new guidelines for knowledge representation on the semantic Web [1]. A thorough review of the semantic Web field comprising of two decades of research has been carried out by [2–4]. A review of semantic Web technologies and its applications in the construction industry, biographies, enhancing the functionality of augmented reality, etc., was presented by [5–7]. Knowledge extraction for the Web of things was designed to automatically identify the important topics from literature ontologies of different IoT application domains by [8]. A systematic review on applications of semantic Web in software testing, distance learning, formal education, and cloud computing was carried out by [9–12].

3 System Design

3.1 Application Architecture

The multi-tier application architecture employed for the implementation of the semantic Web application is depicted in Fig. 1 which depicts the overall application architecture employed in the implementation of semantic Web module detailing out different modules interacting with each other.
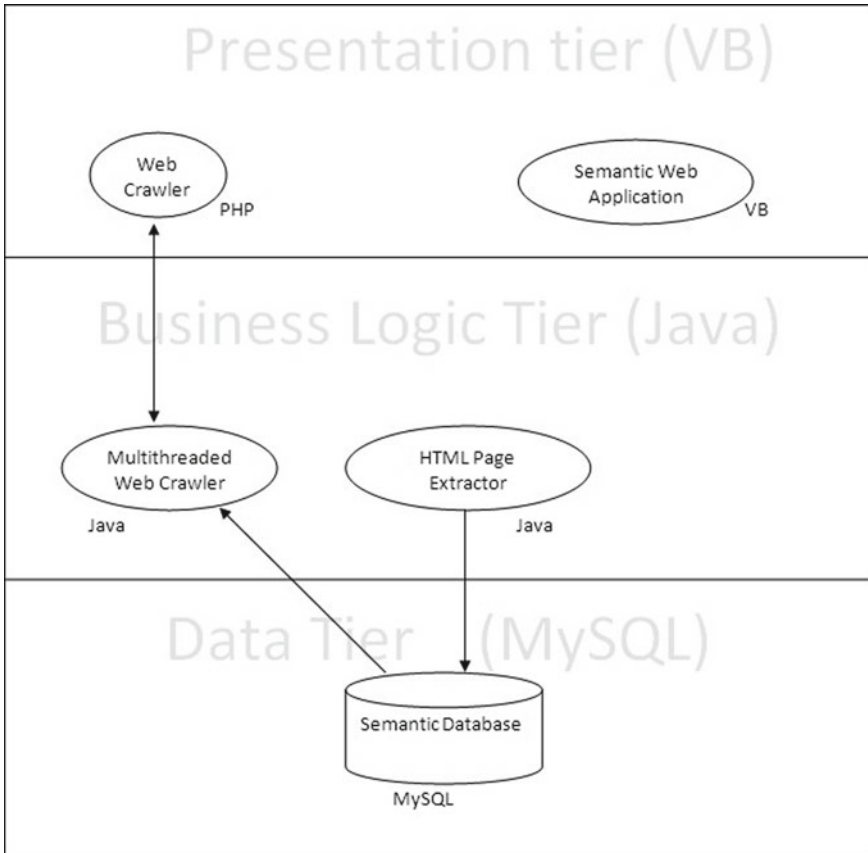
Fig. 1 Semantic Web application framework

3.2   Module Dependency Graph in Presentation Tier

The dependency between different modules of a semantic Web application is shown in Fig. 2.

3.3   Implementation of Multithreaded Web Crawler in Java

Since crawling a Website is an extremely time consuming task, the application becomes too sluggish to respond to the user requests. In order to improve the application's performance, a multithreaded model is employed where task associated with each thread is launching a browser instance in a separate process for execution of a Web crawler. On slow Internet connections, this model experiences a
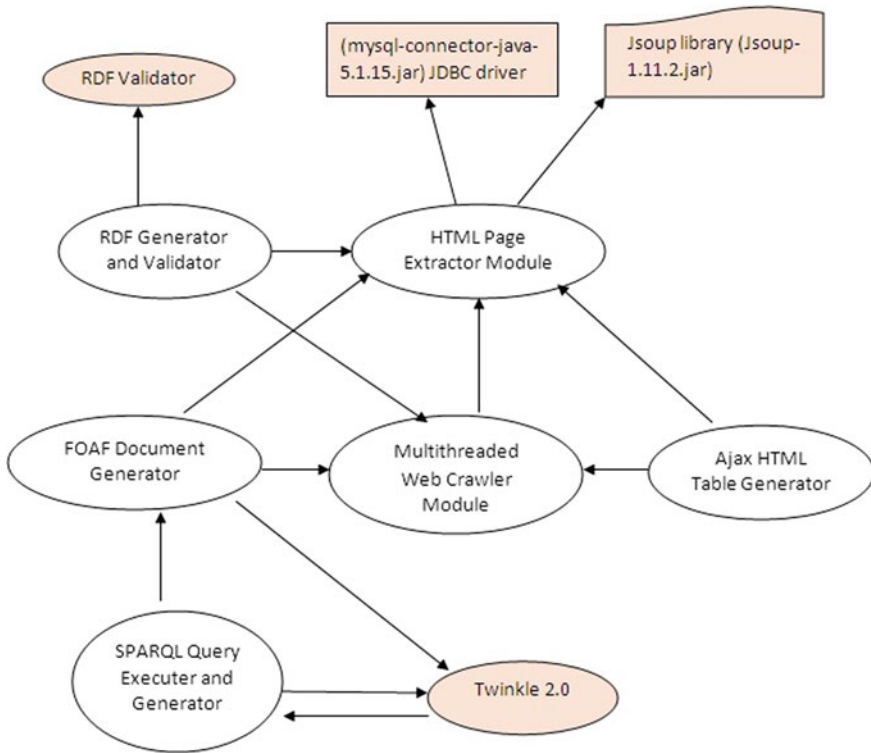
Fig. 2 Dependency between different modules of a semantic Web application

high-performance gain. Figure 3 depicts the working of multithreaded Web crawler module.

The different steps involved in working of Web crawler module are enlisted below:

1. The user generates a message containing subject area of interest.
2. The subject is searched in the MySQL table subject. If the subject already exists in the table, the module queries the user for refreshing the table by pulling data from internet, if the response is 'yes', the module continues its execution, otherwise immediately returns. Also, if the subject does not exist in the table, then the module continues its execution.
3. The HTML page URLs corresponding to the given subject are retrieved from the 'result' table.
4. A separate thread is created for launching a Web browser instance requesting a PHP page for crawling the Website corresponding to the retrieved Web page from the 'result' table.
5. The research attributes pertaining to the journal title, vol no, issues no, impact factor, charges, ISSN. No. of research journal are stored in 'Journal_Details'
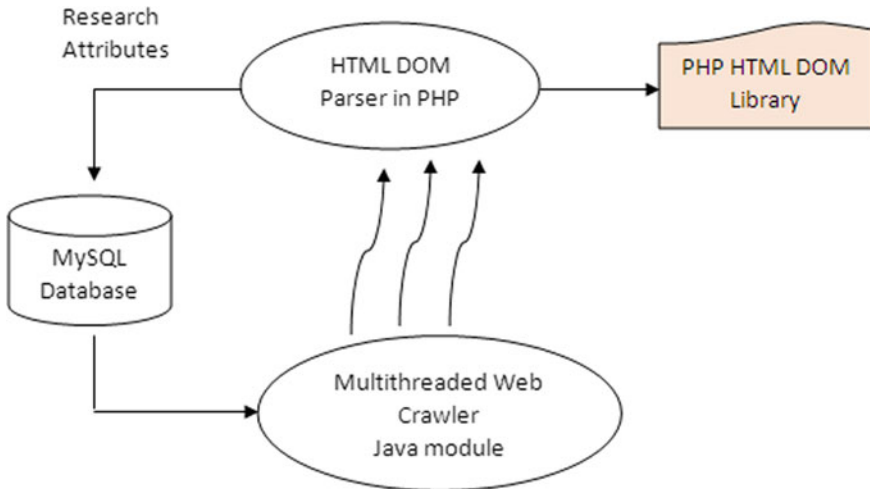
Fig. 3 Working of multithreaded Web crawler module

MySQL table. Index image is downloaded by specifying proper MIME type based on the image file extension.

4   Results and Discussion

An executable batch file with the name 'run_crawler.bat' is created for the automatic execution of the semantic Web module. The batch file is dynamically generated based on the subject selected by an end user and basic operations performed by it are enumerated below:

Working of Semantic Web Module:

- Semantic Web module first looks for queried subject in local MySQL database.
- If the subject exists in database, the requested data is pulled out from the local database.
- If subject does not exist, then query is executed using Google search engine, and relevant HTML pages are retrieved from Internet employing business logic implemented in middle tier using Java technology with the help of JSoup library.
- A multithread application is implemented in Java for parsing each html document by spawning multiple browser instances for invoking html DOM parser implemented in PHP.

Basic operations performed by a batch file, run_crawler.bat are enumerated below:

- Set class path to JSoup java class library, jsoup-1.11.2.jar file
- Set class path to MySQL, Type-IV JDBC driver, MySQL Connector, mysql-connector-

- java-5.1.15-bin.jar.
- Compile Java program, JSoup.java
- Execute Java program, JSoup.java by passing command-line arguments corresponding to Subject ID and Subject Name specified by an end user.

The content of run_crawler.bat executable batch file is shown in Fig. 4.

On execution of the above batch file the Web crawler actin kicks in which spawns multiple threads for searching the entire Web for research journals in the specified subject area. The no. of threads to be created and maximum no. of journals to be retrieved are configurable by an end user. The background processing of multithreaded Web crawler is depicted in Fig. 5.

The output generated by multithreaded Web crawler module is fed as an input to the RDF module which generated RDF triplet for storing the data in machine readable format. The RDF document generated is consumed by HTML Table Generator AJAX
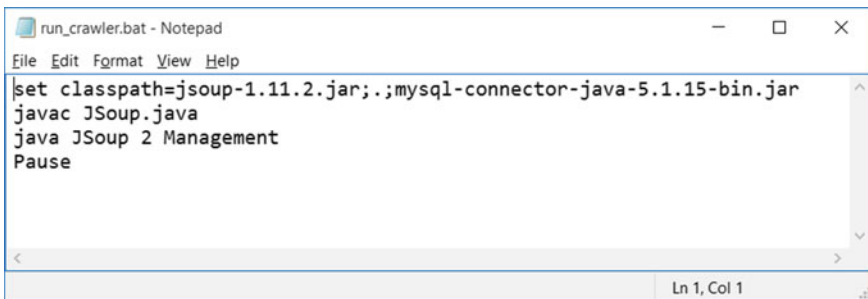
Fig. 4 Content of run_crawler.bat executable batch file
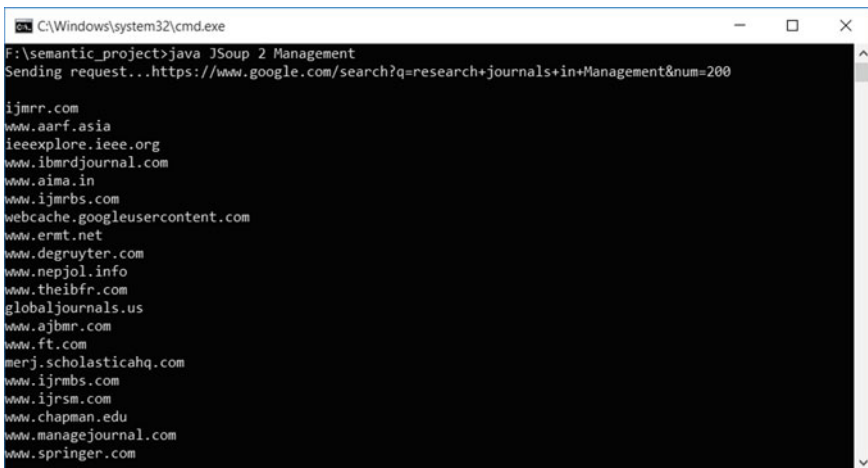
Fig. 5 Background execution of multithreaded Web crawler

Fig. 6 HTML table generated by 'HTML Table Generator' AJAX module

module which displays research journal attributed in the selected subject area such as journal title, volume, issue, processing charges, impact factor, *e*-ISSN, and *p*-ISSN. Figure 6 depicts the HTML table generated by 'HTML Table Generator' AJAX module.

5   Conclusion and Scope for Future Work

The semantic Web renders the search, machine friendly rather than human friendly. In the current research, authors have designed a model for semantic framework which interfaces with different modules employing sematic technologies for retrieving the Web data in machine readable format. The query optimization is achieved through multithreaded crawler which spawns multiple browser instances and retrieves data attributes of interest employing HTML DOM API. For its operation, multithreaded Web crawler interfaces with HTML page extractor module which employs JSoup API for extracting HTML pages residing in the domain from the Web by firing a Google search query. The output generated by multithreaded Web crawler module is fed as an input to the RDF module which generated RDF triplet for storing the data in machine readable format which is consumed by HTML Table Generator AJAX module for displaying presentable data to an end user. The model can be improved further for incorporating thread optimization based on available bandwidth and machine configuration.

References

1. Bonatti PA et al. (2019) Knowledge graphs: new directions for knowledge representation on the semantic web (dagstuhl seminar 18371). Dagstuhl Rep 8
2. Hitzler P (2021) A review of the semantic web field. Commun ACM 64(2):76–83
3. Patel A, Jain S (2021) Present and future of semantic web technologies: a research statement. Int J Comput Appl 43(5):413–422
4. Seeliger A, Pfaff M, Krcmar H (2019) Semantic web technologies for explainable machine learning models: a literature review. In: PROFILES/SEMEX@ ISWC 2465, pp 1–16
5. Sobhkhiz S et al. (2021) Utilization of semantic web technologies to improve BIM-LCA applications. Autom Constr 130:103842
6. Hyvönen E et al. (2019) BiographySampo–publishing and enriching biographies on the semantic web for digital humanities research. In: European semantic web conference. Springer, Cham
7. Lampropoulos G, Keramopoulos E, Diamantaras K (2020) Enhancing the functionality of augmented reality using deep learning, semantic web and knowledge graphs: a review. Vis Inf 4(1):32–42
8. Noura M et al. (2019) Automatic knowledge extraction to build semantic web of things applications. IEEE Internet Things J 6(5):8447–8454
9. Taher KI et al. (2021) Efficiency of semantic web implementation on cloud computing: a review. Qubahan Acad J 1(3):1–9
10. Dadkhah M, Araban S, Paydar S (2020) A systematic literature review on semantic web enabled software testing. J Syst Softw 162:110485
11. Bashir F, Warraich NF (2020) Systematic literature review of semantic web for distance learning. Interact Learn Environ 1–17
12. Jensen J (2019) A systematic literature review of the use of semantic web technologies in formal education. Br J Edu Technol 50(2):505–517

Medical Image Classification Based on Optimal Feature Selection Using a Deep Learning Model

M. Venkata Dasu, R. Navani, S. Pravallika, T. Mahaboob Shareef, and S. Mohammad

Abstract Medical images can be used to diagnose and treat disease. It provides valuable information about internal organs for treatment and assists physicians in diagnosing and treating disease. We use deep learning improved classifier for the classification of lung cancer, brain tumors, and Alzheimer's disease in this study. The classification of medical images is a serious problem in computer-based diagnoses. Hence, we have created a classification of medical images based on selection of the optimal feature using the deep learning model by adding specific steps such as preprocessing analysis, feature selection, and classification. The most important purpose of the study was to find the most accurate selection model for medical image classification.

Keywords Classification · Deep learning · Medical image · Features · Crow search algorithm · Optimization

1 Introduction

Many investigations in the subject of "medical image inquiry" have been undertaken for diagnostic and clinical trials purposes. When it comes to image classification problems, the descriptiveness and discriminative strength of the retrieved features are both critical for improved categorized results [1]. The machine learning component is essential because it is used in many applications in data mining, forecasting models, media information recovery, and other domains. Medical picture databases are used for picture categorization and education. It usually includes drawings that illustrate a variety of characteristics in various circumstances, as well as detailed explanations [2]. When it comes to disease detection in medical diagnostics, it's vital to recognize the most significant risk factor.

M. Venkata Dasu (✉) · R. Navani · S. Pravallika · T. Mahaboob Shareef · S. Mohammad
Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: dassmarri@gmail.com

When a large number of features are deferred by the feature extraction method, manual feature selection becomes necessary. To computerize this method, the appropriate feature selection is explained. Several studies on lung cancer image categorization have been conducted. The computational complexity is relatively significant, and these methods necessitate detailed knowledge of the image structure [3]. Most of the division algorithms reveal a strong connection between an element of an object and its membership function. In addition, accurate identification of its core function is required. The deep neural network (DNN), the proposed method, has been used successfully in real-world segregation methods such as speech recognition, error detection, medical diagnosis, and so on.

It is an underlying machine understanding platform that uses a few statistics to try to express the abnormal state of the current data by using an in-depth layout with a few layers with the potential for direct and indirect switching. One of the most time-saving aspects of the machine learning technique is the reduction in the requirement for feature development, which is one of the advantages of deep learning models [4]. When dealing with large amounts of data and visuals, it can quickly recognize and interpret spoken language, solve problems, and work more efficiently. It is feasible to prevent frequent problems that take up a lot of time by employing deep learning algorithms [5].

Objective

The primary goal of this approach is to establish whether the condition is malignant or benign. This method's operation is primarily dependent on picture classification. This classification is carried out by utilizing the ANN technique in conjunction with the KNN classifier algorithm, which allows us to quickly characterize the disease. Gray Level Co-occurrence Matrix (GLCM) and Gray Level Run Length Matrix can be used to extract features from the input (GLRLM). The qualities of these matrices are used as features to help the classifier get better results.

2   Methodology

2.1   Existing Method

Convolution Neural Network (CNN)

Another type of deep learning neural network is the convolutional neural network (CNN). CNN is a huge improvement over image recognition. They are widely used to analyze visual images and are often involved in separating background images. CNN has three layers: input layer, output layer, and hidden layers. Hidden layers include convolutional layers, ReLU layers, integration layers, and fully integrated layers [6].

Inputs are subject to conversion function using convolution layers. The information is then passed on to the next section. In the next layer, coagulation combines the output of multiple clusters of neurons into a single neuron. Every neuron in a single layer is connected to each neuron in the next layer of fully connected layers. Only a small portion of the pre-layer input is acquired by neurons in the dynamic layer. Each neuron in a fully connected layer receives information from all the elements of the previous layer. CNN is a machine that removes features from images. You no longer need to remove features in person. Features not yet fixed! They are taken as the network operates through a collection of images. In-depth learning models are particularly accurate in computer vision systems because of this. Feature detection is taught on CNNs in more than a dozen or hundreds of hidden layers [7]. The complexity of the features studied grows with each layer. The result is a modified feature map. The size of the output image is smaller than the actual image.

Support Vector Machine (SVM)

Support vector machine (SVM) is classified as classification algorithm; however, they can be used to solve classification issues. It is capable of dealing with both continuous and categorical variables. To differentiate various classes, SVM creates a hyperplane in multidimensional space. SVM creates the optimum hyperplane, which is then used to minimize an error. SVM's purpose is to divide a dataset into different classes. To divide or classify two classes, SVM creates a decision boundary, which is a hyperplane between them. SVM is also utilized in picture classification and object detection.

Disadvantages

- The accuracy of the data is determined by its quality.
- If there is a lot of data, the prediction stage may take a long time.
- Conscious of the data's size and irrelevant aspects.
- High memory is required since all of the training data must be stored.

2.2 Proposed Method

When raw data is fed into an input image, it is preprocessed, and the result of that preprocessing is fed into the feature extraction. It extracts features based on the GLCM and GLRLM characteristics. These features are saved, and CSO is used to pick features for the input. After picking the features, the TREE classification is used to determine whether the output is dangerous or incorrect, and accuracy can be used as a performance measure. The three stages of this medical imaging phase are preliminary analysis, feature selection, and classification. Figure 1 shows the paradigm of the medical image classification diagram.

Pre-processing: The goal of pre-image processing is to improve image quality. Histogram scaling is used to enhance image quality input in this case.
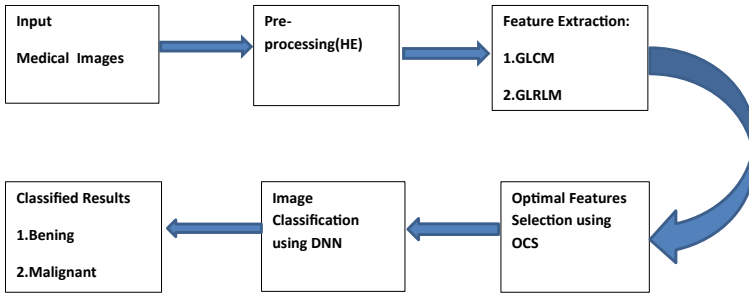
Fig. 1 Block diagram of proposed methodology

Optimal feature selection (OCS) is a method of determining which attributes are most significant to you. Feature selection is a crucial image processing technique that facilitates image classification.

Classification: It divides medical images into benign and malignant categories based on the best features selected.

ANN

Artificial neural network (ANN) is a learning program that uses connected nodes or neurons in a horizontal structure to replicate the human brain. The neural network may be trained to recognize patterns, separate data, and predict future events by studying data. The input to a neural network is broken down into levels of abstraction. It, like the human brain, can be trained to recognize patterns in voice or images using a variety of examples. Its behavior is determined by the way its many pieces are connected, as well as the strength, or weights, of those connections. These weights are automatically changed during training using a learning algorithm until the artificial neural network successfully completes the task.

ANN-Based Classifiers: ANN-based classifiers are commonly utilized in pattern recognition applications for classification. For a specific problem, ANN will be trained to behave as a classifier. The linear perceptron classifier, feed-forward networks, radial basis networks, recurrent networks, and learning vector quantization (LVQ) networks are all examples of this construction. The training and testing phases of an ANN-based categorization system work together. The labeled data are utilized to train (learn) the network and develop the classifier model during training. During testing, unlabeled data is sent into the trained network (model), which is then categorized using the model created during the training phase.

Learning via *Decision Trees*: The decision tree method is common method which is used in data mining. The aim is to develop a model that can forecast the value of a target variable using a set of input variables. A decision tree is a straightforward approach of categorizing cases. Assume that all of the input characteristics have finite discrete domains and that only one goal feature, "classification," exists in this section. A class is allocated to each element of the categorization domain. A decision

tree, also known as a classification tree, is a tree with an input characteristic labelled on each internal (non-leaf) node.

The arcs that arise from a junction named after an input feature are either named after each potential value of the target feature or directed to a subordinate decision node named after a different input feature. By training each fresh instance to emphasize previously mis-modeled training examples, boosted trees gradually create an ensemble. AdaBoost is an excellent example. These can be applied to situations involving regression and classification. A random forest classifier is a random forest-based bootstrap aggregate.

Decision trees are a form of data mining technique that combines mathematical and computational tools to assist in the analysis of large amounts of data.

3 Results

The proposed methodologies are implemented through the use of the MATLAB computer- aided design program. The version of the tool used in this research work is of R2018a with required communication and image processing toolboxes.

In this, when the input image is given it undergoes histogram equalization of the input image and then some features get extracted and gives resultant image which is categorized as either benign or malignant and displays its accuracy. (Figs. 2, 3, 4 and 5).

Fig. 2 Input image
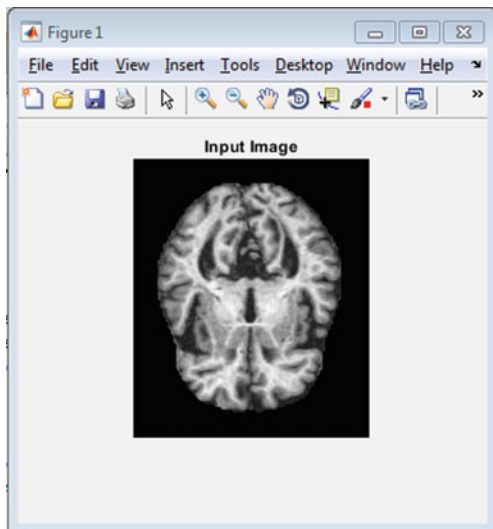
Fig. 3 Histogram equalization of input image
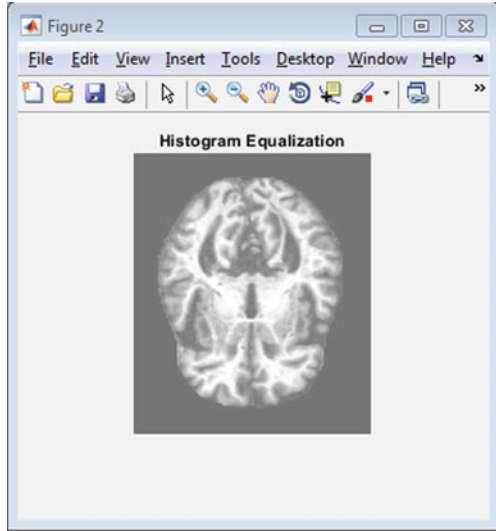
Fig. 4 Classification result
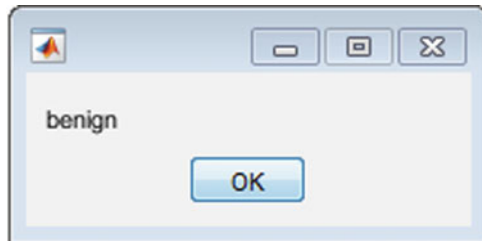
Fig. 5 Accuracy
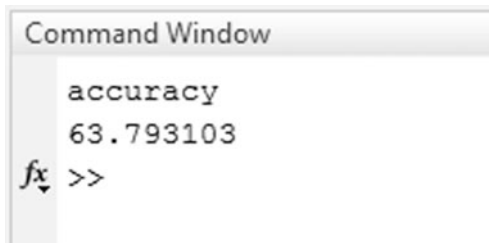
4 Conclusion

A novel classification strategy was proposed to identify medical photographs, which chose the best attributes from the images. As a result, a one-of-a-kind soft set-based clinical picture classification system was developed in order to achieve superior accuracy, precision, and computation speed. Other quantitative elements could be added to the existing techniques to increase the consistency of the classification of

complex medical images. The method used in this case was the most costly. In the future studies, segmentation systems and certain automatic classification methods, as well as a feature reduction strategy, should be employed to detect the tumor component in medical images.

References

1. Bharath R, Mishra PK, Rajalakshmi P (2018) The automatic measurement of fatty tissue formation using curvelet transform and SVD. Biocybern Biomed Eng 38(1):145–157
2. Park SB, Lee JW, Kim SK (2004) Content-based image classification using the neural network. Pattern Recogn Lett 25(3):287–300
3. CH Nagaraju, AK Sharma, MV Subramanyam (2018) Reduction of PAPR in MIMO-OFDM using adaptive SLM and PTS technique. Int J Pure Appl Math 118(17):355–373
4. Sayed GI, Hassanien AE, Azar AT (2019) The selection of a feature in a chaotic novel crow search algorithm. Neural Comput Appl 31(1):171–188
5. Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. SpringerPlus 5:507. https://doi.org/10.1186/s40064-016-2152-2
6. Bauer S, Wiest R, Nolte LP, Reyes M (2013) MRI-based medical imaging study analysis of brain tumor studies. Phys Med Biol 58(13):1–44. https://doi.org/10.1088/0031-9155/58/13/R97
7. Zhang QL, Zhao D, Chi XB (2017) Review for in-depth readings based on medical imaging diagnoses. Comput Sci 44:1–7

Real-Time Indian Sign Language Recognition Using Image Fusion

Tejaswini Kurre, Tejasvi Katta, Sai Abhinivesh Burla, and N. Neelima

Abstract According to a 2018 report by the World Health Organization, nearly 63 million people in India suffer from partial or complete hearing impairment. They communicate using sign language (SL), a language that uses a set of specific non-verbal gestures, mainly involving hands. One major barrier in their way of communication is that the majority of the non-hearing and speech impaired population does not understand sign language, which raises the necessity of developing sign language recognition systems that can be standardized across the nation. This paper aims at giving the best recognition model for Indian Sign Language. The system involves the Bag of Visual Words technique for real-time prediction of the ISL with a comparative study of feature detectors and descriptors like SIFT, SURF, ORB, STAR + BRIEF, and FAST + FREAK. The results indicate that using the SURF feature detector and descriptor along with SVM yields an accuracy of 99.94% and gives a rotation threshold of up to 5 degrees. Meanwhile, parallel experimentation and comparison with the concept of CNN and ASL dataset have shown promising results such as 100% accuracy for ISL with CNN, 65.41% accuracy for ASL with SURF and SVM, and 99.77% accuracy for ASL with CNN.

Keywords Bag of visual words · Feature extraction · Speeded up robust features · Support vector machine · Classification

1 Introduction

According to WHO, more than 5% of the global population suffer from moderate to severe forms of deafness, mutism, and hard hearing [1]. Since most people have no prior knowledge or experience with sign language, they take the help of sign language interpreters. Technology makes communication easier between conventional and impaired people by eliminating sign language interpreters. Sign language

T. Kurre (✉) · T. Katta · S. A. Burla · N. Neelima
Department of Electronics and Communication Engineering, Amrita School of Engineering, Bengaluru, India
e-mail: tejaswinikuree@gmail.com

(SL) consists of sign/gestures created with the hands and various other bodily movements, together with facial expressions, used fundamentally by deaf and dumb people to communicate. ASL and most other sign languages use one hand, while ISL uses both hands for gesture representation, making the interpretation complex, owing to which there is significantly less research and development on this subject [2].

The paper attempts to propose a fast, efficient, and accurate ISL recognition model based on the comparison among various feature detection and extraction algorithms such as SURF, SIFT, ORB, STAR + BRIEF, and FAST + FREAK. The system acquires vision-based image data through the webcam of user computer, which is then processed using the feature detection and extraction algorithms, and the classifiers such as SVM, KNN, LR, and NB are used to classify data according to the predefined categories. The translation is in the form of a gesture to English text.

2 Literature Survey

Shravani et al. [2] proposed a model that uses image pre-processing techniques and Bag of Visual Words model for extracting features from the input dataset, following which histograms are generated. At the last step, the obtained features are fed into the BOVW model for the classification.

Tripathi et al. [3] uses a key frame extraction methodology to extract the gesture frames and obtain the features using orientation histogram with Principal Component Analysis (PCA).

Rokade and Jadav [4] system approach involves segmenting the hand based on skin color statistics, then converting that segmented image into binary, and applying feature extraction on the binary image; for extraction of the features, the techniques used are distance transformation, Discrete Fourier Transform, and probability distribution property that is central moments; and for classification and recognition, Artificial Neural Network and SVM are used.

Neelima and Sreenivasa Reddy [5] has proposed a model which consists of three steps: pre-processing, feature extraction, and classification. The pre-processing stage includes skin masking and histogram matching. In the feature extraction stage, Eigenvalues and Eigenvectors were considered, and to recognize the sign, Eigenvalue-weighted Euclidean distance is used.

A vision-based sign recognition model using feature fusion methodology which involves histogram of oriented gradients, shape descriptors [6] and SIFT involving both the hands, and the experimental analysis demonstrated that the fusion model gives better results than other existing methods. The advantage of Support Vector Machine as classifier is described in [7].

Ansari and Harit [8] utilizes the Microsoft Kinect console to obtain the image data instead of common cameras, which gives the advantage to use Kinect's infrared sensor to obtain depth information in addition to RGB color information. Image depth information can improve the performance of a practical sign language translator but can be unsuitable for real-life applications, involving common cameras.

3 Proposed Methodology

The proposed methodology consists of 6 major steps as shown in Fig. 1, namely vision-based data acquisition, image pre-processing, feature extraction, codebook construction by clustering, classification using various classifiers under Bag of Visual Words model, and gesture recognition. Images are acquired through webcam, and the redundant information is removed using image pre-processing. The Bag of Visual Words model is implemented on the processed images for detecting features and extracting their descriptors using feature detection and description algorithms SURF, SIFT, ORB, STAR + BRIEF, and FAST + FREAK. The obtained descriptors are passed on to clustering algorithm (mini-batch K-means), and histograms are generated for each image. Classification is done using prominent machine learning algorithms such as SVM, LR, KNN, and NB.

3.1 *Image Pre-psrocessing*

Images are pre-processed to perform feature extraction and description. Pre-processing is done in 3 steps as shown in Fig. 2, namely skin masking, skin detection, and Canny edge detection. First, the raw image is converted to HSV and gray color spaces, and the skin pixels are extracted from the HSV image to create a skin mask. By using skin masking, the hand in the gray image is extracted. At last, the Canny edge detection technique is used to detect sharp edges in the image, thus showing the edges of the hand in the image. This helps in reducing the redundant information as the edges of the hand are sufficient for hand gesture recognition.
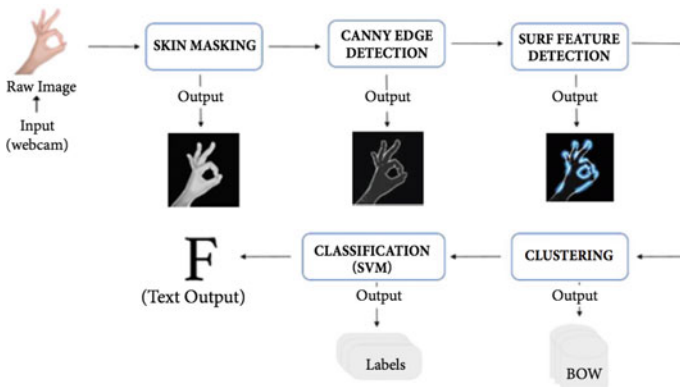
Fig. 1 Block diagram of the proposed methodology
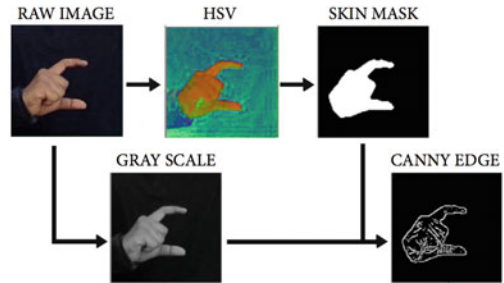
3.2 Feature Extraction

The Bag of Visual Words model is used to represent each image as a frequency histogram of feature clusters. This is achieved in 3 steps—feature detection and description, visual vocabulary development, and frequency analysis. First, features and their descriptors for each pre-processed image are acquired using one of the algorithms-SURF, SIFT, ORB, STAR + BRIEF, and FAST + FREAK, and a comparative study for all algorithms is done. The feature descriptors obtained are clustered using mini-batch K-means algorithm to construct a visual vocabulary. Then, clustering is done to cluster all similar features into 'K' number of clusters. The value of K used is 280 as there are 35 classes (35 * 8 clusters). The frequency histogram for each image is then calculated using visual words technique. The output of Bag of Visual Words is a feature vector whose elements are the frequencies of each cluster.

3.3 Classification

The penultimate step is to classify the processed images into one of the 35 classes using machine learning algorithms.

The dataset is divided into training and testing in the ratio 80:20. The model is trained and tested for different machine learning classifiers including Support Vector Machine, K-Nearest Neighbors, Logistic Regression, and Naïve Bayes for comparison on how the system works for different algorithms.

3.4 Gesture Recognition

The last step is to develop real-time recognition model based on previous steps. SURF algorithm for feature detection and description and SVM algorithm for gesture recognition are used as this combination has the highest accuracy. Images are acquired in run-time video stream captured by a web camera. The acquired image

is pre-processed, progressed to the Bag of Visual Words model to obtain frequency histogram, and the sign language is detected using SVM in real time.

4 Experimental Results and Analysis

To determine the suitable combination of the decision-making pipeline with the proposed approach, different classification algorithms with various feature detection and extraction techniques are considered. SURF produces substantially accurate predictions for the considered classifiers among all feature detection and extraction algorithms. Another notable observation is that K-Nearest Neighbor classifier shows the least accuracies for $K = 3$ as per the proposed model, which is unsuitable for large dimensional data and quite unstable due to the trial-and-error approach to determine K value, giving the least testing performance. This is shown in Fig. 3.

Figure 4 delineates the performance of the model proposed in terms of the time taken by mini-batch K-means clustering for every feature detector and descriptor combination.

It can be observed that SURF is almost 4.5 faster in comparison with SIFT, which can be attributed to its use of integral image and box filter. SURF uses only a 64-D descriptor vector as compared to the 128-D descriptor vector of SIFT. This results in fast feature computation and matching capability. BRIEF descriptors with STAR feature detector outperform all the other techniques as it represents the image patch as a binary string which makes it efficient to compute and store in memory. FREAK descriptors with FAST key point detectors also yield fast results, while ORB which uses FAST detectors and BRIEF descriptors with additional features exhibit
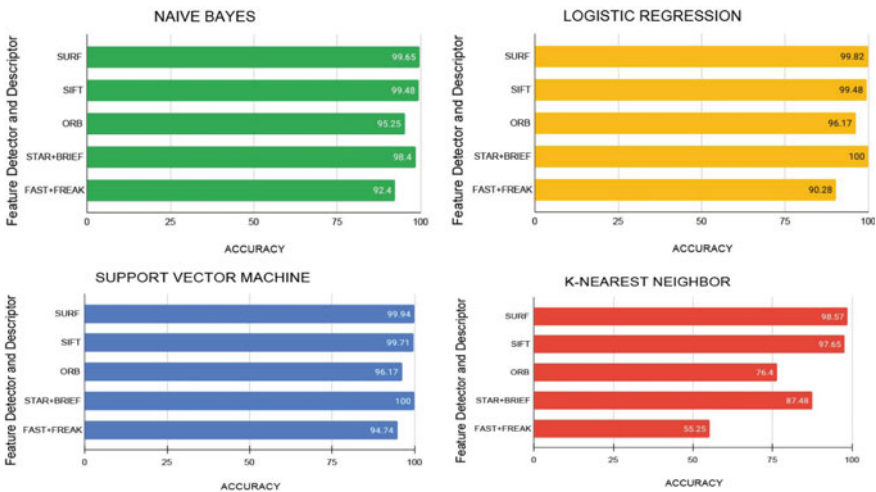
Fig. 3 Comparison between various feature detectors and classifiers
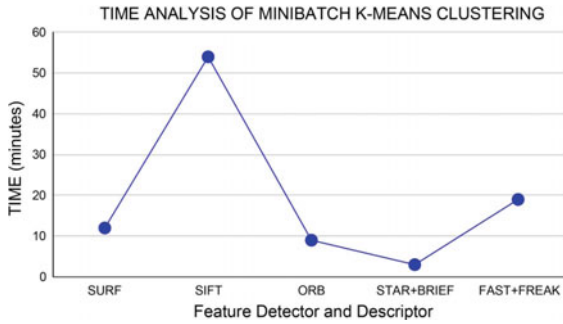
Fig. 4 Performance of mini-batch *K*-means w.r.t feature detectors
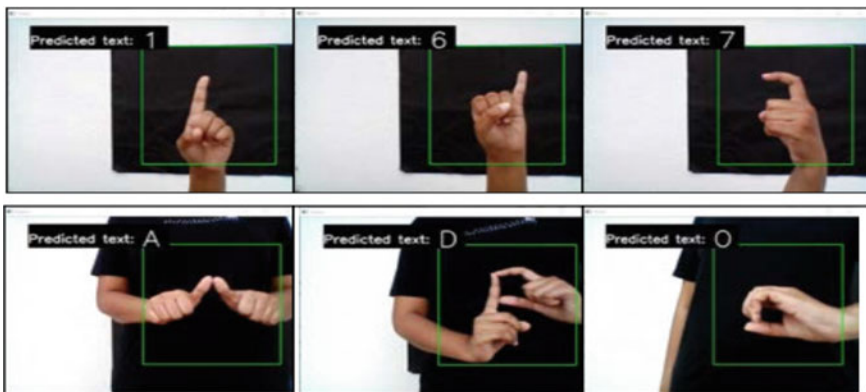
Fig. 5 Output alphabets (*A*, *D*, *O*) and output numerals (1, 6, 7)

satisfactory performance and low computational cost when compared to SIFT and SURF. The sample output detections were shown in Fig. 5.

The real-time recognition model, involving the BoVW technique, integrated with robust SURF feature descriptors and, using an SVM classifier, scored a 99.94% accuracy. The precision, recall, and *f*1 score recorded w.r.t. to micro-averaging were 99.94%. The presented paper also provides a comparative study on the resulting accuracies, when the input dataset is changed ISL to ASL, as shown in Fig. 6.

5 Conclusion

This paper proposes a real-time model for ISL gesture recognition, based on the incoming video stream in the form of sequences of images, acquired from the web camera. The proposed technique revolves around the Bag of Visual Words model, which involves feature extraction and detection using SURF, SIFT, ORB, FAST,

Fig. 6 Comparison between various classifiers for the SURF feature descriptor using ASL
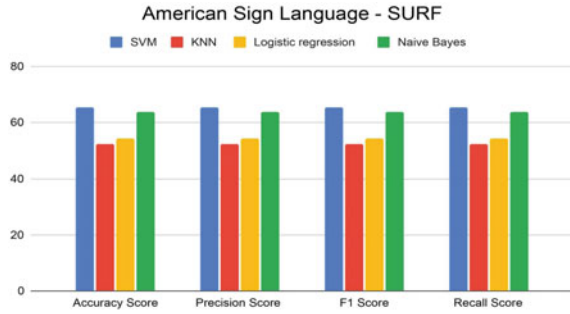
and BRIEF algorithms, construction of visual vocabularies by mini-batch *K*-means clustering, followed by frequency analysis through histogram creation, and classification/prediction using prominent classifiers like SVM, KNN, Naïve Bayes, and Logistic Regression. The model achieves the maximum accuracy of 99.94% with the SVM classifier implemented with the SURF algorithm for the real-time model. The model achieves an accuracy of 65.41% when the ISL dataset is replaced with an ASL dataset, as the ASL dataset considered for this model has great variation among the individual images. The paper can be further extended by using deep learning techniques like CNN for feature extraction and detection.

References

1. Rajeswari N, Priyadharsini N (2017) Perlustration of deaf-and-dumb alphabet detection and interpretation. Int J Eng Dev Res (IJEDR) 5(2). e-ISSN: 2321-9939
2. Shravani K, Sree Lakshmi A, Sri Geethika M, Kulkarni SB (2020) Indian sign language character recognition. IOSR J Comput Eng (IOSR-JCE) Ser I 22(3):14–19. e-ISSN: 2278-0661, p-ISSN: 2278-8727
3. Tripathi K, Baranwal N, Nandi GC (2015) Continuous Indian sign language gesture recognition and sentence formation. In: Eleventh international multi-conference on information processing—2015
4. Rokade YI, Jadav PM (2017) Indian Sign language recognition system. Int J Eng Technol (IJET) 9(3S)
5. Neelima N, Sreenivasa Reddy E (2017) An efficient approach to CBIR using DWT and quantized histogram. Int J Innovative Comput Inf Control IJICIC (Scopus-Q1) 13(1):157–166
6. Neelima N, Sreenivasa Reddy E (2016) An integrated approach to CBIR using multiple features and HSV histogram. Int J Eng Technol (IJET) 8(5)
7. Neelima N, Sreenivasa Reddy E (2015) An efficient multi object image retrieval system using multiple features and SVM. In: Advances in signal processing and intelligent recognition systems. Springer (Scopus), pp 257–265
8. Ansari ZA, Harit G (2016) Nearest neighbour classification of Indian sign language gestures using Kinect camera. Sadhana 41:161–182
9. Karami E, Prasad S, Shehata M (2017) Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images. In: Newfoundland electrical and computer engineering conference

Design IoT-Based Smart Agriculture to Reduce Vegetable Waste by Computer Vision and Machine Learning

Himanshu Pal and Sweta Tripathi

Abstract We all know that as the population grows, there will be a shortage of vegetables due to restricted farming land supplies. If land resources are limited and output is limited, the expanding population's needs will not be met. As a result, we've allocated our resources in such a way that they can meet the demands of an expanding population. With the inclusion of smart farming, there is a need to design a system that can evaluate veggies and notify farmers about their state on a regular basis, so that farmers are aware of their exact status and may take appropriate action. As a result, farmers will be able to predict when to harvest, resulting in reduced vegetation loss. In this paper, we have designed a system that uses computer vision and machine learning algorithm to detect the real-time condition of vegetables over time and take the appropriate steps to prevent vegetable waste.

Keywords Computer vision · Drone in agriculture · IoT · Machine learning · Decision tree · Smart agriculture

1 Introduction

We now realize that food safety is the most pressing issue on the planet. According to a research issued by the World Health Organization (WHO) in 2017, around 600 million individuals are impacted by eating unfit food [1]. According to that research, around 420,000 persons were died in one year as a result of eating contaminated food. These factors make it difficult to evaluate food quality and determine food conditions in order to preserve human health. There are many methods by which we can monitors fruits and vegetables. Traditional and visible ways of examining the physical composition of meals, such as texture, sensitivity, taste, freshness, and color, produce ambiguous results and risk infecting the food by contacting the

H. Pal (✉)
Electronics and Communication Engineering, Kanpur Institute of Technology, Kanpur, India
e-mail: himanshu13feb@yahoo.com

S. Tripathi
Head EC Department, Kanpur Institute of Technology, Kanpur, India

specimens [2]. Another method detects microorganisms and counts their quantity in food samples; based on the idea that food spoils owing to the development of hazardous microbes, especially perishable foods like fish and meat. Methods which are explained above are detected by directly contacting with them, which would be contaminating that produce. Therefore, there is a need of system which will detect contaminated fruits and vegetables without direct contact with them. Today, we can observe that the world's population is steadily expanding. We also recognize that resources are finite, and we must make the best use of our limited resources to ensure that every person on the planet has enough food. If there is a food crisis, agriculture will be unable to meet the needs of the entire world's population. Hence, agriculture must be intelligent.

Farming is becoming smarter by the day. However, there are still a few vulnerabilities that haven't been addressed and the resources are not fully utilized. There is still a room to make the agriculture systems more efficient and intelligent. We notice that a greater proportion of fruits and vegetables are still wasted in the field due to a lack of sufficient care. Farmers are frequently unaware of when their fruits and vegetables are ready. We all know that fruits and vegetables spoil when they are not harvested on time or are not properly cared.

We have worked on a smart system that can quickly determine the state of vegetables. This technology will give the farmer feedback on the veggies so that they can pick them and take appropriate measures before them decay. We can boost the yield of vegetables in this way. When productivity rises, it helps to meet the demands of an expanding population.

We explain in the research gap how fruits and vegetables decay as a result of improper maintenance and cultivation delays. In the issue statement, we will explain how we can conserve our fruits and veggies in some way. This entire study is focused on tomatoes where a method is developed to conserve tomatoes, lowering the percentage of rotting tomatoes. We're utilizing a drone which is equipped with multispectral camera [3] to assess the condition of the tomatoes. We can lower the proportion of rotting tomatoes by employing this technology since we receive correct feedback from the smart system and, as a result, we can harvest our produce without further delay.

2 Literature Review

This section covers research on the food evaluation system, which is part of the smart agriculture process. Here, we are reviewing various study articles on similar themes, and as a result, we have discovered some scope where some work has to be done in order to smarten the agricultural system. The author of [4] examines the imbalance in supply chain management between the industrial and agricultural sectors, as well as the challenges with vendor-specific production systems. The authors of [5] focus on

a smart agriculture application that uses computer vision and ultrasonic emission to create virtual barriers to shield crops from ungulate assaults, resulting in considerable reductions in production losses. In [6], author introduces a drone-based system to detect cattle's using SSD 500 and YOLO-V3 and CNN.

In [7], author provides a holistic approach in using different deep neural networks (DNNs) object detector methods including YOLO-V3 to create bird detection models using aerial photographs captured by unmanned aerial vehicle (UAV). In [8], author uses the Raspberry Pi and demonstrated his work by using devices to repel birds in optimizing crop production in agriculture.

The author of [9] provides a thorough examination of the sensors, deployment, and analytical methods used in water quality monitoring. From a technological standpoint, the validity, consistency, and security of traceability information are guaranteed in [10] author study on a food safety traceability system based on big data and the Internet of things. This is a good way to improve the credibility of traceability data, ensure data integrity, and optimize the data storage structure.

In [11], the author focuses on using more than two reference measures, such as evapotranspiration and moisture, under various situations, which helps to equally partition the connection by the number of sources. The author of [12] seeks to establish a secure measurement and review system based on IoT to monitor the quality of perishables, as well as supply chain management (SCM) with an emphasis on transportation, without the need for human interaction. Based on the monitoring findings, the author develops a unique deep neural network model in [13] to forecast multiple states of quality of food. The research suggests detecting total volatile organic compounds (TVOCs) inside food packaging, which have been produced during food spoilage, to monitor variations in food quality.

To categorize date fruits, author in [14] employed machine learning and a vision system. Using the VGG-16 and AlexNet architectures, the system presented in [14] was utilized to classify the kind and ripeness of date fruits. It has a five-stage maturity accuracy of 97.25% and a date type classification accuracy of 99.01%. ML-based approaches were employed in a number of studies to classify the maturity of Milano and Chonto tomatoes by author in [15] as well as Philippine coconut by author in [16]. The study [16] used the SVM classifier (SVM) and random forest methods to divide the Philippine coconut into three stages of development (over-mature, mature, and premature). SVM and Bayesian-ANN were used by Nyalala et al. [17] to predict the weight and quantity of cherry tomatoes related to depth pictures from 2 and 3D photographs.

Author in [18] classifies maturity level of tomato, varieties of tomatoes by using neural network and yielding an accuracy of 99.31%. Using morphological traits and RGB, HSI, and YIQ spaces, the author of [19] describes an algorithm enabling robots to detect the locate of ripe tomatoes. Their precision was 96.36%; however, their robot failed to detect and locate obstructed tomatoes. The image segmentation for fruit recognition and yield estimate was achieved using MLP and CNN by the author in [20]. Nonetheless, there were segmentation mistakes in areas where the

image quality was low owing to bad lighting. To distinguish ripe tomatoes, the author in [21] employed laboratory color space and luminance in-phase quadrature phase (YIQ) color space with wavelet transformation picture fusion characteristic.

3 Methodology

3.1 *Computer Vision*

Computer vision is an interdisciplinary scientific topic concerned with how computer models may be built to get high-level comprehension from digital pictures or movies in order to develop autonomous systems similar to the human visual system [22]. Combining machine learning or deep learning with computer vision has allowed computers to better grasp what they see, which has aided computer vision advancements. Computer vision has been created to ease the process of detection, pattern recognition, and prediction. It can automatically extract complicated characteristics that are not built by technologists and can do so by learning from various training data. Computer vision mainly consists of two major parts, that is, hardware and software. Hardware part contains vision device that is camera and source of light. Software parts gather images and analyze the images and further extract information for which we are using this technology. Computer vision consists of major three steps: (1) acquiring images and videos, (2) processing those images and videos, (3) understanding and extracting information from images and videos. Figure 1 shows the fundamental steps which are essential for the information extracted from images and videos.

Fig. 1 Fundamental steps for computer vision
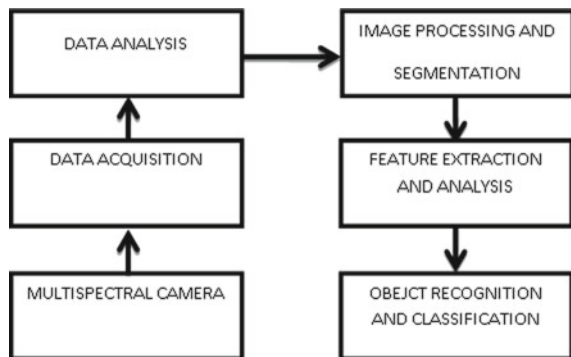
3.2 Multispectral Camera

A multispectral camera is widely used in agriculture [23]. If we go through RGB camera basically it has three color bands that are red, green, and blue. But in case of multispectral camera, it detects 3 and 5 bands of spectrum (as shown in Fig. 2). In case of 3-band multispectral camera the detection color spectral bands are blue, green, red, and near-infrared. In the case of 5-band multispectral camera, the detection color spectrum bands are blue, green, red, and red edge and near-infrared band.

The near-infrared band is beyond the red band. We cannot see this band by our eyes but it can detect the condition of plants very accurately. We know that healthy plants reflect more near-infrared light than an RGB camera. When we walk through sick plants, we see a reduction in near-infrared reflectance.

Another spectral region of relevance is the red band. This band is located between both the red and near-infrared bands. The reflection of plants increases between both the red and near-infrared regions, resultant a substantial rise in reflection across the red edge area as in Fig. 2. The quantity of reflection in this band has been connected to plant health in a significant way.

The quantity of reflection within that band has also been connected to plant health. Due to changes in air conditions, the quantity of reflectance fluctuates from day by day. It is impossible to compare these photographs across time because of this. Furthermore, if the intensity of sunshine varies throughout a flight, some portions of a field look darker or brighter than others. A down-welling light sensor is included in certain passive multispectral cameras to compare measured reflectance values from picture to image. For each of a camera's spectral bands, DLS measures the quantity of sunlight as in Fig. 3.

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

where NIR and red stand for the reflection in the near-infrared and the red spectral bands.
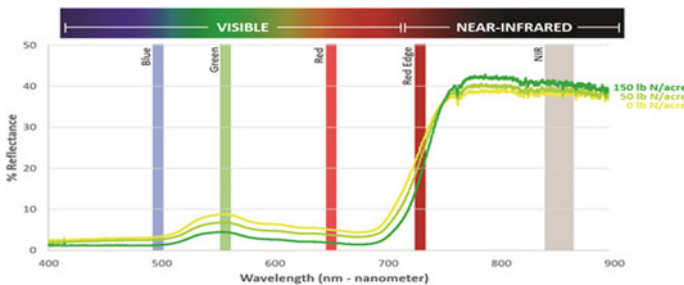
Fig. 2 Invisible and near-infrared spectral reflection of healthy and strained plants [23]
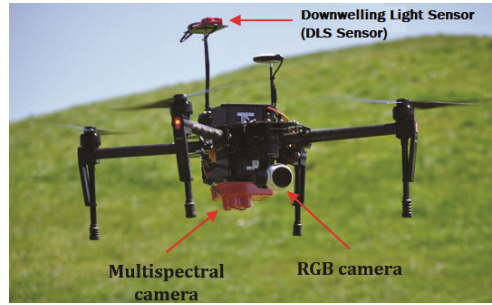
Fig. 3 A quadcopter featuring a DLS sensor, an RGB camera (DJI), and a multispectral camera (MicaSense RedEdge) (MicaSense) (Image courtesy of MicaSense, https://www.micasense.com/kits/.)

NDVI is greatest at portraying variations in chlorophyll concentration and canopy concentration during the early to mid-periods of the season, but it reaches a maximum value later in the season after canopy closure. Another extensively used VI is the normalized difference red edge (NDRE), which is calculated as:

$$\text{NDRE} = \frac{\text{NIR} - \text{Red Edge}}{\text{NIR} + \text{Red Edge}}$$

3.3 Software

Software plays an important role in computer vision technology. Data acquisition is the process of selecting the best quality images captured by camera and stores it in database. After acquisition, there is a step for analyzing the images stored in database by using software as follows: R, MATLAB, OpenCV, Minitab-15, Keras, Yolo, etc. Further we have to process those images by using several algorithms, and some of them have been mentioned here such as watershed algorithm, ellipse model, and K-mean clustering.

3.4 OpenCV

OpenCV was created to provide as a foundation for computer vision. This collection contains a large number of machine learning and computer vision algorithms that have been optimized. Object identification, face detection and recognition, and object movement tracking are examples of these algorithms. The programming languages C++, Python, Java, and MATLAB are all supported by OpenCV. Read and write

pictures, store and capture photographs/videos filter and transfer images, and crop detection are all standard features in OpenCV.

3.5 Classification of Images by Using Machine Learning Algorithm

Image classification is a crucial job in computer vision since it is used to recognize an item in a picture. This task requires determining the likelihood of a given visual item class appearing in incoming pictures. In addition, the ultimate objective of computer vision is to develop machine learning or deep learning models that accurately mimic differentiating sample attributes. The relationship between the precise measured characteristics of a sample and its spectral information was used to develop a machine learning and deep learning model.

The calibration or training set, and the validation or prediction set, are usually included in the samples used to build the model. The calibration collection consists of a few typical specimens that are used to calculate the model's parameters. Support vector machine (SVM) is common machine learning algorithm and likewise deep learning modeling approach artificial neural network (ANN), convolutional neural network (CNN), and the convolutional neural network (CNN).

Decision Tree

A decision tree is a visual depiction of all the different options for making a choice. It is a method of sorting a large number of observations into multiple categories. We are analyzing the data based on a few conditions and then dividing it into several categories. To run some predictive analysis, we categorize it, and this categorization method answers queries such as whether this data falls in category 1 or category 2. A common machine learning method, the decision tree, is an effective tool for categorization issues.

A multistage or hierarchical organizational decision scheme or a tree-like architecture underpins the decision tree. Internal nodes and terminal nodes, also known as leaf nodes, make up the tree. Each internal node has a decision function that indicates which node should be visited next, whereas each terminal node displays the result of a specific input vector that leads to this node. Each node in the decision tree-like structure makes a binary choice that separates one or more categories from the rest of the groups. Moving down the tree until you reach the leaf node is how most processing is done. Binary recursive segmentation, which is an iterative procedure for splitting data into partitions, is used to build a decision tree.

4   Results and Discussion

4.1   Problem Statement

We are aware that several studies are being conducted to boost vegetable output. After reading a number of study articles on various ways for increasing the production of vegetables and fruits, we discovered that there is still a lot of research to be done in order to boost vegetable yield. We've seen veggies go bad most of the time when they're grown in the field due to farmers' ignorance. Most farmers are unaware of the best times to plant vegetables; therefore, we're creating a system with the use of computer vision, multispectral camera, and machine learning algorithms to address this issue. This will identify crop health, sickness, and maturity of vegetables, among other things. This will identify crop health, disease, and vegetable maturity and provide information to farmers so that they may take the right action.

4.2   Designing

Data Collection

We gathered raw data from many sources, such as the Internet, and captured photographs by visiting to fields in my town and organizations to create our machine learning system. We collect data related to tomatoes with these things keeping in mind (1) Pictures of tomatoes at various stages (2) Pictures of both healthy and unhealthily ripe tomatoes (3) Pictures of rotten tomatoes (4) Pictures of healthy and diseased tomato plants (5) Pictures of curl leaves.

Different Stages of Tomatoes

From Fig. 4, we observe that a tomato goes through six different stages of color change: green, tannish yellow, combination of tannish yellow, pink, and red, combination of pink and red, orange, and red. For the ease of simplicity here, we overlap stages 2–3 and 4–5 and this overlapping stage we denote by yellow and orange respectively. Now we consider four stages, where a green color tomato denotes immaturity, a yellow tomato denotes a breaker, an orange tomato denotes the pre-harvest stage, and a red tomato denotes the harvest stage.

Preparation of Data

We filtered our raw data and divided it into groups based on its attributes and color after gathering data from numerous sources. We construct a group of infected tomatoes, a group of healthy tomatoes based on their phases, a group of healthy plants and sick plants, and a group of curl leaf-related photographs. Outliers in the gathered data are identified and removed. The model will not perform well if the data is not adequately specified.
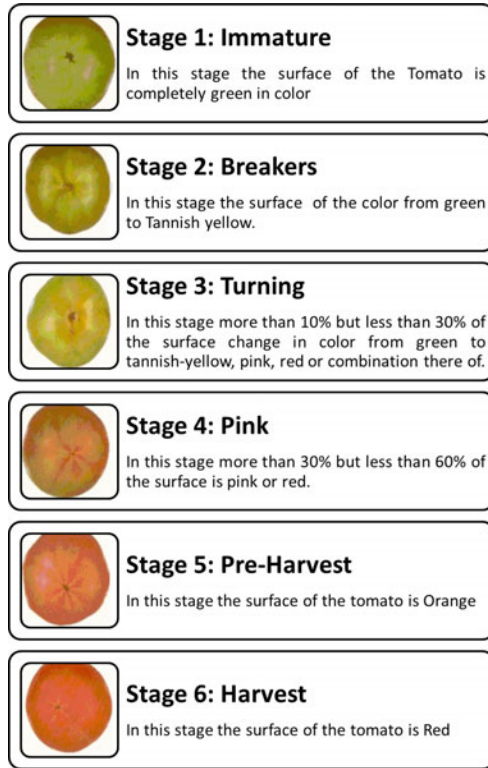
Fig. 4 Classification of maturity of tomato

Data Modeling

Now that we've gathered our data from various sources, we've divided it into groups based on their characteristics. Now we've trained our model to match my problem statement; the training process is depicted in Fig. 5. We utilized the random forest machine learning technique in our model. We detect distinct phases of tomatoes, tomato health, and disease that will affect tomato development using a random forest algorithm. In this case, we'll use two separate decision trees: one will recognize different phases of tomatoes as shown in Fig. 7, and the other will detect the health of plants and vegetables in the field that are impacted by various illnesses as shown in Fig. 6. Finally, the output of both trees is concatenated. Finally, the outputs of both trees are integrated to build a model known as a random forest machine learning method.
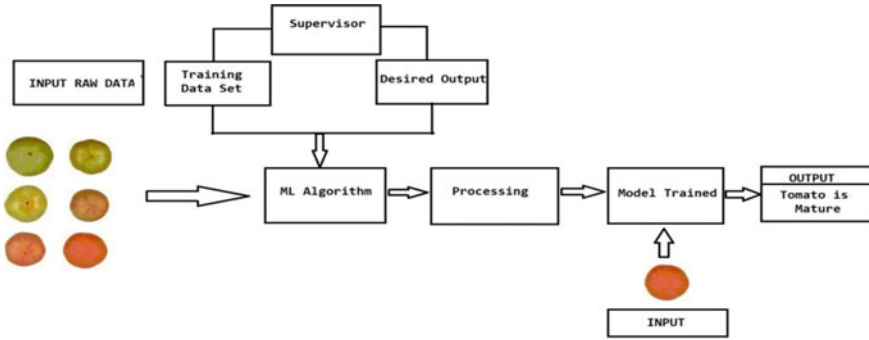
Fig. 5 Supervised machine learning model
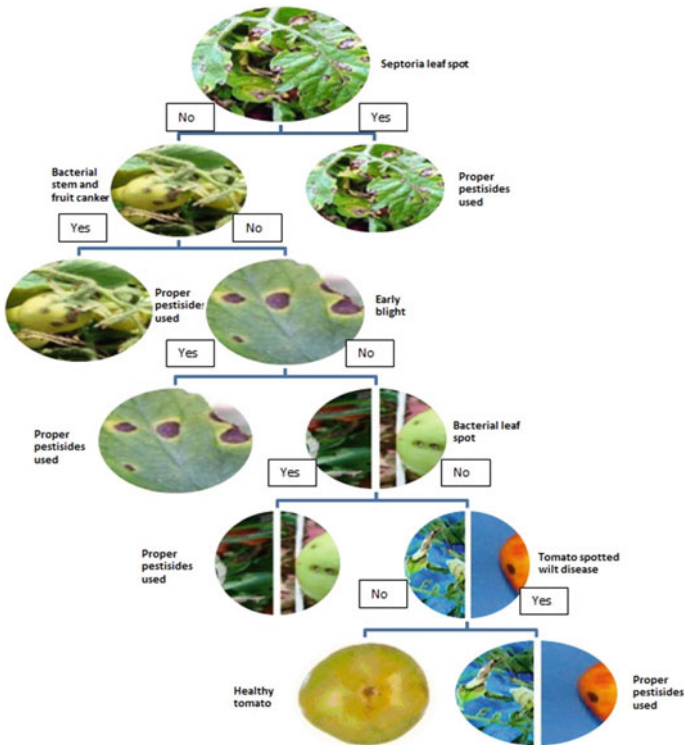
Fig. 6 Decision tree for tomato disease detection
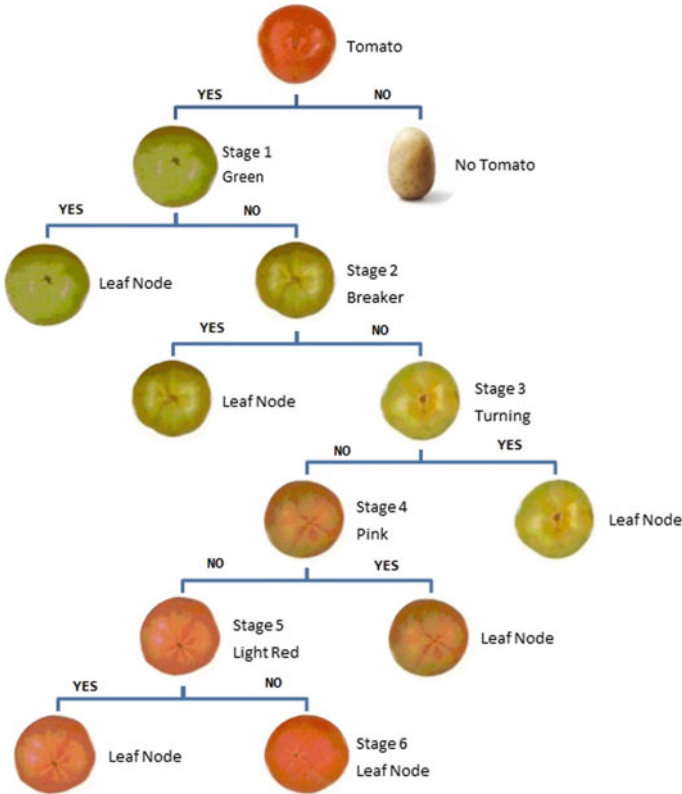
Fig. 7 Decision tree for tomato maturity detection

4.3 Results

In this section, we are going to discuss the outcome which we are getting by our experiment. Our research aims to measure the condition of tomatoes time to time and gives the result accordingly, our research main goal is to increase the productivity of tomato. Somehow we are successful in my experiment; our system is monitoring the crop from germination to till harvest. We programmed our drone in such a way that it will adjust according to the dimension of field. It will follow the path as shown in Fig. 8. We indicate each line with some identity so that we come to know the yield in that indicated location.

To begin, all of the tomato training samples are utilized to identify the tree's structure. The method then splits the data into two parts using every feasible binary split and chooses the one that minimizes the total of the squared deviations from the mean in the two parts. We trained our machine learning model before further processing. We know that decision tree is a supervised type of machine learning algorithm so there is a need of training before implementing into the system.
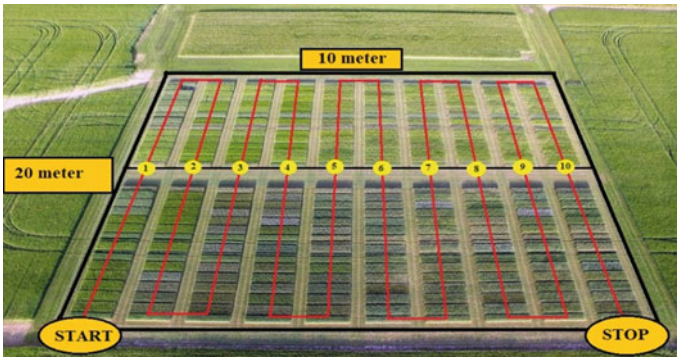
Fig. 8 Movement of drone on the field

When model got trained we now apply decision tree algorithm for the separation of the kinds of tomatoes. Here our main aim is to reduce the waste of tomatoes while growing. When system detects the stage 2 that is breaker stage, we pluck the tomatoes from the field as per the location detected by system because this will keep the quality of the tomatoes' taste and flavor. Furthermore, your tomatoes will be protected against fruit cracking, sunscald, and blossom end rot as a result of this. It will also aid in the ripening of the other immature fruits on the vine.

After that, the splitting procedure is done to each of the newly created branches. This procedure is repeated until each node achieves the minimum node size chosen by the user and becomes a terminating node and drone follow the path on the field as shown in Fig. 8.

Decision tree algorithm basically segregates all color tomatoes, and here we are mainly focused on tomatoes which are green in color, we pick tomatoes in this stage from the fields, it will ripe naturally after few days, and we get more time for the tomatoes to supply in market. In this way, we somehow reduce the wastage of tomatoes and increase the yields of tomatoes and also we come to know the productive of tomatoes in field.

The production of tomatoes and infected areas in specified locations is shown in Figs. 9 and 10, respectively. This graph describes all the possible outcomes while harvesting tomatoes; accordingly, we have to pick tomatoes from the field.

5 Conclusion

The use of Internet of things technology to control food quality and safety can help to prevent serious food safety events from occurring. Furthermore, the problem may be recognized in real time and with greater accuracy, and the source of the danger can be swiftly identified, ensuring the food's quality. We all know that resources are

Fig. 9 Production of tomatoes in specified location
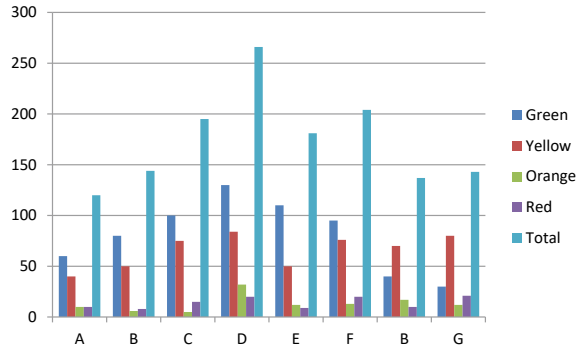
Fig. 10 Bar graph of infected area

	1	2	3	4	5	6	7	8	9	10
Septoria Leaf Spot	1	0	0	0	0	0	0	0	1	1
Bacterial Stem and fruit canker	0	0	0	0	0	0	0	0	0	1
Early Bligth	1	0	0	0	0	0	0	0	0	0
Bacterial leaf spot	0	0	0	0	0	0	0	1	1	1
Tomato spotted wilt disease	1	1	1	1	0	0	0	0	0	0
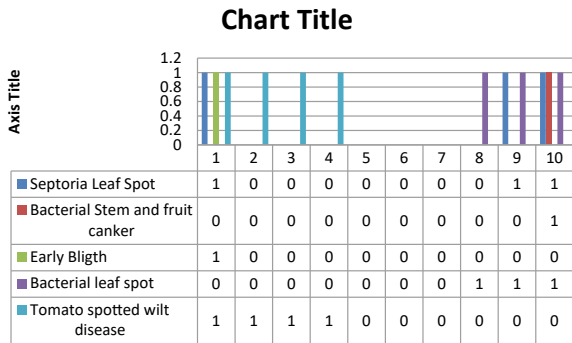
finite, and the world's population is growing by the day. If resources are not used efficiently, they will not be able to meet the needs of every individual on the world.

So, keeping this in mind, we're here to build a tomato-growing system that will boost tomato output while simultaneously improving tomato quality. We are employing a computer vision machine learning algorithm and multispectral camera in this experiment, which has the potential to identify crop quality. The color of the tomatoes will be detected by this camera. We discuss four stages of tomatoes in this paper: immature, breaker, pre-harvest, and harvest. Tomatoes are green when they are immature, yellow when they are breaker, orange when they are pre-harvest, and red when they are harvested. When a tomato is in the harvest stage, it is at risk of rotting if it is not plucked off the vine. Farmers, according to reports, harvest their tomatoes ahead of schedule to avoid rotting.

References

1. WHO. Food Safety. Retrieved from https://www.who.int/news-room/fact-sheets/detail/food-safety. Accessed 31 Oct 2017
2. Cheng JH, Sun DW, Han Z, Zeng XA (2014) Texture and structure measurements and analyses

for evaluation of fish and fillet freshness quality: a review. Compr Rev Food Sci Food Saf 13(1):52–61

3. Laben CA, Brower BV (2000) Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. U.S. Patent 6011875, 4 Jan 2000

4. Almadani B, Mostafa SM (2021) IIoT based multimodal communication model for agriculture and agro-industries. IEEE Access 9:10070–10088

5. Adami D, Ojo OM, Giordano S (2021) Design, development and evaluation of an intelligent animal repelling system for crop protection based on embedded edge-AI. IEEE Access 9:97444–97456

6. Aburasain RY, Edirisinghe EA, Albatay A (2021) Drone-based cattle detection using deep neural networks. In: Arai K, Kapoor S, Bhatia R (eds) Intelligent systems and applications. Springer, Cham, pp 598611

7. Hong SJ, Han Y, Kim SY, Lee AY, Kim G (2019) Application of deep-learning methods to bird detection using unmanned aerial vehicle imagery. Sensors 19:1651

8. Roihan A, Hasanudin M, and Sunandar E (2020) Evaluation methods of bird repellent devices in optimizing crop production in agriculture. J Phys Conf Ser 1477:032012

9. Manjakkal L, Mitra S, Petillot YR, Shutler J, Scott EM, Willander M, Dahiya R (2021) Connected sensors, innovative sensor deployment, and intelligent data analysis for online water quality monitoring. IEEE Access 8:13805–1824

10. Zheng M, Zhang S, Zhang Y, Hu B (2021) Construct food safety traceability system for people's health under the internet of things and big data. IEEE Access 9:70571–70583

11. Alghazzawi D, Bamasaq O, Bhatia S, Kumar A, Dadheech P, Albeshri A (2021) Congestion control in cognitive IoT-based WSN network for smart agriculture. IEEE Access 9:151401–151420

12. Nasir M, Bhutta M, Ahmad M (2021) Secure identification, traceability and real-time tracking of agricultural food supply during transportation using internet of things. IEEE Access 9:65660–65675

13. Lam MB, Nguyen TH, Chung WY (2020) Deep learning-based food quality estimation using radio frequency-powered sensor mote. IEEE Access 8:88360–88371

14. Altaheri H, Alsulaiman M, Muhammad G (2019) Date fruit classification for robotic harvesting in a natural environment using deep learning. IEEE Access 7:117115–117133

15. Pacheco WDN, Lopez FRJ (2019) Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K-NN, MLP, and K-means clustering. In: Proceedings of the 22nd symposium on image, signal processing and artificial vision (STSIVA), pp 1–5

16. Caladcad JA, Cabahug S, Catamco MR, Villaceran PE, Cosgafa L, Cabizares KN, Hermosilla M, Piedad EJ (2020) Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal. Comput Electron Agric 172. Art. No. 105327

17. Nyalala I, Okinda C, Nyalala L, Makange N, Chao Q, Chao L, Yousaf K, Chen K (2019) Tomato volume and mass estimation using computer vision and machine learning algorithms: cherry tomato model. J Food Eng 263:288–298

18. Wan P, Toudeshki A, Tan H, Ehsani R (2018) A methodology for fresh tomato maturity detection using computer vision. Comput Electron Agric 146:43–50

19. Are A, Motlagh AM, Mollazade K, Teimourlou RF (2011) Recognition and localization of ripen tomato based on machine vision. Aust J Crop Sci 5(10):1144–1149

20. Bargoti S, Underwood JP (2017) Image segmentation for fruit detection and yield estimation in apple orchards. J Field Robot 34:1039–1060

21. Zhao Y, Gong L, Huang Y, Liu C (2016) Robust tomato recognition for robotic harvesting using feature images fusion. Sensors 16(2):1–12

22. Choi I, Kim J, Jang J (2018) Development of marker-free night-vision displacement sensor system by using image convex hull optimization. Sensors 18(12):4151. https://doi.org/10.3390/s18124151

23. Michał M, Wiśniewski A, McMillan J (2016) Clarity from above: PwC global report on the commercial applications of drone technology. PwC Drone Powered Solutions. Retrieved from https://www.pwc.pl/pl/pdf/clarity-from-above-pwc.pdf

Analytical Study of Hybrid Features and Classifiers for Cattle Identification

Amanpreet Kaur, Munish Kumar, and Manish Kumar Jindal

Abstract Animals biometric identification system is emerging in pattern recognition and machine learning due to its diversification of applications and its uses nowadays. This new standard has received more attention due to its biometric features for identification of cattle. In this article, the authors have observed three feature extraction techniques that are Scale Invariant Feature Transform (SIFT), Speeded up Robust Feature (SURF) and Oriented Fast and rotated BRIEF (ORB) from the collected dataset. The classifiers, i.e., decision tree, k-NN, and random forest identify breed of cattle based on efficacy of extracted features. Experimental results are conducted on a dataset with accuracy of 97.23%.

Keywords SIFT · SURF · ORB · Decision tree · k-NN · Random forest

1 Introduction

Cattle identification systems are allured nowadays to identify cattle using the image by which we can find out which class cattle are associated with. Such a system is required for its intensive applications as experts in the field can easily recognize their breed without defining any specific parameters using biometric traits. Muzzle point in animals is the primary feature investigated as a distinguished pattern since 1921 [6]. Factors required for a cattle identification system are cattle traceability to track the natural habitat of cattle, health trajectory to record the infectious diseases, vaccination record and insurance claims in banks. For this purpose, we require such a

A. Kaur (✉)
Department of Computer Science and Applications, Guru Nanak College, Sri Muktsar Sahib, Punjab, India
e-mail: amanpreetkkour07@gmail.com

M. Kumar
Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda, Punjab, India

M. K. Jindal
Department of Computer Science and Applications, Panjab University Regional Centre, Sri Muktsar Sahib, Punjab, India

model which involves a new paradigm of machine learning algorithms that train data as a sample of images from the database. The machine is trained by extracting various features that are important parameters to affect the overall efficacy of the system. In this study, authors have experimented using some of feature extraction techniques, feature selection, and classification algorithm in order to promote a methodical and reliable identification system for recognition of cattle breed. The current approach is implemented with local binary pattern (LBP) to draw out local invariant features from the nasal area of cattle, consist of distinct feature. The experiments are implemented on the dataset of 930 images generated by authors of their own. Dataset for experiments is partitioned into a training phase that has 80% of the total dataset and the rest 20% is into the testing phase. Furthermore, the authors have also performed three-fold cross-validation dataset partitioning. This paper comprises of total six sections. Section 1, is a brief discussion about the problem. In Sect. 2, we have presented the related literature survey. In Sect. 3, data collection is presented for the proposed study. In Sect. 4, is about the methodology, its design is detailed with all the components and methodology of the work done. In Sect. 5, the authors detailed the experimental results of the study. Finally, Sect. 6, summarizes the work done so far.

2　Literature Survey

A related literature review in the field of cattle identification using distinct biometric features is reviewed. Literature survey assesses the studies for different feature extraction techniques and classification for individual cattle and breed of cattle where muzzle pattern is used as primary feature for identification. This study also compared the experimental results achieved by the authors in terms of recognition accuracy. Awad et al. proposed muzzle point as primary biometric feature for cattle identification. The experiment evaluated with 93.3% accuracy [1]. Tharwat et al. the authors examined the local binary pattern (LBP) technique for feature extraction and linear discriminant analysis (LDA) feature reduction with 99% accuracy [11]. Tharwat et al. the authors of the paper explored the Gabor feature to extract the features from grayscale images. Two-level fusion, i.e., feature and classifier, is done to get a more independent feature vector with 99.5% of identification accuracy [12]. Tharwat et al. the authors proposed the biometric model for cattle identification system based on the Gabor feature extraction technique. The authors experimented on the dataset of 31 and achieved a 99% identification rate [10].

El et al. the author proposed the bovine's classification of the muzzle using k-NN classifier and artificial neural network classifier. The author used the dataset of 28 bovines for their experiments. The result evaluated that KNN gives 100% accuracy, whereas ANN produces accuracy of 92.76% [4]. Gaber et al. have also presented similar work based on Weber's Local Descriptor (WLD) for background removal. The author achieved identification accuracy of 99.5% from images of cattle [3].

Table 1 Feature extraction and classification techniques used by authors

S. No.	Feature extraction	Feature classification	Dataset	Accuracy (%)	References
1	Local invariant features using SIFT	RANSAC algorithm for matching	6 × 15 = 90 images	93.3	[1]
2	Weber local descriptor and adabooster	k-NN, Fk-NN	1200 images	99.5	[3]
3	Gabor filter based and feature fusion	SVM classification using kernels and classifier fusion	NA	95.5	[12]
4	Watershed technique	Haar-based classifier	431 subjects	95	[5]
5	Weber local descriptor, local binary pattern, feature fusion	SVM Decision tree k-NN	900 images	96.5	[8]

Kusakunniran et al. examined Haar feature-based cascade classifier to train the system to identify cattle from the images. Watershed technique is tried to segment the muzzle area of cattle (Region of Interest). These experimental results have 95% accuracy on the dataset of 431 subjects [5]. Sian et al. explored their research work to identify individual cattle from the group of single breeds. The texture feature is extracted with Weber's local Descriptor and Local binary pattern techniques and feature fusion obtained from both. Results reveal that SVM has the best recognition result than decision and k-NN, i.e., 95.3% [8]. The literature survey is summarized in Table 1 that includes feature extraction, classification, dataset, and accuracy of their work.

3 Data Collection and Pre-processing

Authors have collected a database using the camera with features like large screen, wider lens decent number of megapixels-16 megapixels, and telephoto camera. The database is collected from the distinct primary sources. The dataset was collected from distinct primary sources like from the farm of village Chak-Khrung, District Fazilka, Punjab, India of breed (1) Holstein Friesian (HF1) breeds with 470 images. (2) Jersey 200 images (3) Rathi 100 images from Punjab progressive dairy farmer association, Jagraon, Punjab, India. (4) And, the database of Sahiwal breed from the Northfields farm, Sri Muktsar Sahib, Punjab, India of 160 images. Dataset consists of 930 images collected from distinct sources of 4 breeds. Dataset initial level was manual segmented from face to muzzle point. The dataset consists of all types of images such as (1) Illuminated images, (2) Degraded images, (3) Unclean images,

and (4) Partial images as a part of our database. The accuracy achieved by this system is degraded as a dataset is composed of blurred and distorted images.

4 Methodology and Design of Proposed System

The proposed cattle identification system works into these phases:

(1) Initially, Data processing and Pre-processing phase, here we gathered the images from primary and reliable sources. All the collected data is normalized to 64 × 64-bit pixels.

(2) Second Feature extraction phase is the most deciding phase which affects the system efficacy. Numerous feature extraction techniques do exist that help to recognize an image of cattle, few of which are SIFT, SURF, ORB, FAST, BRISK, BRIEF. This study is confined to SIFT, SURF, and ORB feature extraction algorithm.

(3) Next determining phase is feature selection that is used to reduce the computations by removing the irrelevant feature. Feature selection also improves the accuracy of the system and reduces complexity by using supervised or unsupervised techniques. While using these SURF, SIFT, and ORB algorithms, heap of memory space is needed to store feature vector. K-means clustering algorithm partition the data into clusters and locality preserving projection (LPP) is a reduction algorithm to reduce the size of the feature vector by considering unique features.

(4) Classification phase is to predict if the input images are labeled accurately or not. This prediction depends on the stored images of the training. The efficacy of the cattle identification system entirely depends upon quality of features extracted from the database. Then cattle images are classified using decision tree, random Forest, and k-NN.

4.1 Feature Extraction Techniques

Feature extraction is prime phase for image recognition because the system performance mainly depends upon type and quality of features being extracted from image. This information extracted from Region of Interest to increase the system performance after the initial phase of preprocessing. For this purpose, feature extraction techniques namely SURF, SIFT, and ORB extract some discriminative features for identification like grooves, valleys, ridges, and bead structures from cattle muzzle patterns will be considered for experiments.

Scale Invariant Feature Transform (SIFT). SIFT descriptor extensively used in object and pattern recognition to extract features presented by Lowe [6]. We give an input image to SIFT to extract distinct features and it defines an output for the

same. SIFT algorithm describes local features of any digital captured image. SIFT is a patent algorithm and also invariant to geometric transformation. SIFT generates many features for even small objects and they can be extended. SIFT works into phases, in the first phase i.e., scale-space; location for finding features using Gaussian kernel computed with DoG function to identify the most promising features invariant to scale and orientation. DoG analyze the image at different scales. In the second phase of key point localization to locate accurate key points using DoG. Key point's location and scale are identified based on the stability. In the third phase, orientation is assigned to the key points with orientation histogram. In the fourth phase, key point descriptors are the furnished points as a high-dimensional vector by 16×16 window used to detect object from an image. In the last phase, key point matching where two images are matched by identifying its nearest neighbor. In some cases, accurate results are not produced on blur and illuminated images. SIFT descriptor is better than SURF in scaled images) [6]. SIFT cannot detect the images containing uniform and silent features [9].

Speeded up Robust Feature (SURF). SURF is presented by Bay et al. [2] to extract local features from an image for object recognition and is a patented algorithm. SURF works by applying a Gaussian second derivative mask to an image at many scales. SURF algorithm has three sections; Interest points detection using hessian matrix in which image is converted into integral image. Interest points are described in two steps:

1. Orientation assignment using Haar wavelet
2. Descriptor components by constructing a square region around key points, then split into 4×4 subregions is laid over interest points. Square wavelet response is computed from 5×5 samples and then sum up the subregions of length 64 bits. Then the descriptor components obtained from images are compared and matching pairs are found. Concept of integral images and box type convolution filters makes its faster than SIFT [2]. SURF key points are extracted from the images. SURF cannot detect if the small patterns of images are copied [9].

ORB (Oriented Fast and Rotated BRIEF). ORB, developed by Rublee et al. [7] in the OpenCV lab the algorithm is free to use and has enhanced performance by fusing FAST descriptor and Binary robust independent elementary feature (BRIEF) descriptor. ORB works into phases, in the first phase input image is scaled at different levels. In the second phase extract, Features from accelerated segment test (FAST) features on all levels. Into the third phase apply grid filtering. In the fourth phase extract feature orientation and at last extract feature descriptors. ORB produces multi-scale features, deals with noisy images and rotation invariant [7].

4.2 K-Means Clustering Algorithm

K-means clustering is unsupervised learning used to solve clustering problem of unlabeled dataset into different clusters. Here, K is predefined cluster created in the process. This algorithm performs task to determine the best value for k and then assign each datapoints to the nearest k center and create a cluster based on Euclidean distance. K-means works in the following steps

1. Select value for k.
2. Assign each datapoints to the closest centroids to define predefined clusters.
3. Calculate the variance and place a new centroid for each cluster.
4. If reassignment occurs then repeat step 3 or the model is ready.

4.3 Locality Preserving Projection (LPP)

The purpose of the LPP is dimensionality reduction algorithm to improve features by reducing the random variables and letting the principal variable reside to represent the original data. This method is proposed for unsupervised linear dimensionality reduction LPP perform its work in the following phases

1. Construction of graph Laplacian.
2. Choose the weights on edges.
3. To computing eigenvector equations and eigenvalues, the best is with nonzero value.

4.4 Classification Algorithms

For any identification system, image classification is a prime phase. Image classification classifies images from the extracted features. There is numerous state-of-the-art image classifiers available like support vector machine (SVM) classifier, k-NN classifier, and Naïve Bayes classifier, and random forest classifier, XGBOOSTING, and decision tree classifier. In this study, we are considering decision tree, k-NN, and random forest classifiers.

k-NN Classifier. It's one of the simplest forms of machine learning. k-NN works on the concept to classify a datapoints on the basis how its neighbor is classified. k-NN technique to classify new points is based on similarity measured of previously stored points. If we consider two categories of images, the k-NN algorithm in this case will return the category that has the minimum distance. In k-NN desired data is first loaded then the distance is computed using Euclidean or other methods, sort the computed distance and pick the one with minimum distance.

Decision Tree. This classifier is a tree structured that belongs to supervised learning techniques. In decision trees internal nodes represent features of a dataset that

provides information about class. Branches represent decision rules to make decisions and leaf nodes are the final outcomes of decision nodes. In order to build a tree CART algorithm (Classification and regression tree) is used. It has an advantage as less data cleaning is required in the decision tree than others, less efforts for data preparation during preprocessing.

Random Forest. This is a popular ensembled algorithm in supervised learning technique. Random forest function works in two phases. In the first step, it trains the image dataset so that it creates a set of decision trees by selecting a subset of the training set and then aggregates the final class of test object into a single classifier. The second step, value predicted by all tree and then final value predicted based on majority of votes. Random forest is called multi-class classifier as it has collection of trees that work to classify the data.

5 Experimental Validation and Performance Analysis

This section covers the comparative analysis of work done in the field based on feature extraction techniques and its recognition accuracy attained by the classifiers. To perform the experiments, basic requirements are Intel corei7 processor, minimum 8 GB of RAM and Windows 10 operating system. The programming language is Python 2.3 and the OpenCV image processing library.

The analysis is measured on the basis of classifier wise recognition accuracy. For the experimental work, the dataset composed of 930 images of Sahiwal, Rathi, Jersey and Holstein Friesian (HF1) breeds are captured. In this work, 80% of all collected are used for training set, and the rest 20% of the dataset for testing set.

Table 2 illustrates the classifier wise-recognition accuracy and experimental reports delineating that SIFT (8) + SURF (8) + ORB (8) feature extraction algorithm has achieved best accuracy than others, i.e., 97.23% with random forest classifier. Table 3 depicts threefold cross validation portioning dataset by achieving 93.40% accuracy on SIFT (8) + SURF (8) + ORB (8) considered to be outstanding with random forest classifier than k-NN and decision tree classifier.

6 Conclusion

Cattle identification is an evident field in pattern recognition and computer vision machine learning based by using muzzle point as distinct feature to recognize cattle breed. This research has a dataset of 930 images collected from the farm of Punjab, India of 4 breeds, results are compared to the other dataset using similar types of techniques [1, 3, 8]. The motivation behind of the work is to present SIFT, SURF, and ORB feature detectors and check the efficacy of algorithms to recognize cattle with muzzle patterns. A combination of these descriptors uses a decision tree, k-NN,

Table 2 Classifier wise recognition accuracy for cattle identification

Feature extraction techniques	Classifier wise-recognition accuracy (training: 80% and testing: 20%)		
	k-NN	Decision tree	Random forest
SIFT (8)	76.41	90.76	90.76
SURF (8)	73.82	71.73	78.80
ORB (8)	70.21	87.50	89.67
SIFT (8) + SURF (8)	81.05	91.47	92.56
SIFT (8) + ORB (8)	87.25	92.49	93.36
SURF (8) + ORB (8)	85.19	90.30	91.15
SIFT (8) + SURF (8) + ORB (8)	90.87	94.32	97.23

Table 3 Classifier wise recognition accuracy for cattle identification-threefold cross validation

Feature extraction techniques	Classifier wise—recognition accuracy (training: 80% and testing: 20%) 3-Fold cross-validation dataset partitioning		
	k-NN	Decision tree	Random forest
SIFT (8)	73.94	81.59	82.81
SURF (8)	69.43	62.41	77.76
ORB (8)	71.69	74.02	80.29
SIFT (8) + SURF (8)	75.9	78.97	85.01
SIFT (8) + ORB (8)	83.83	79.87	91.89
SURF (8) + ORB (8)	79.81	79.51	89.39
SIFT (8) + SURF (8) + ORB (8)	86.96	80.94	93.40

and random forest, models. This work can later be expanded on global feature, i.e., contours and shape of muzzle pattern for more accurate results.

References

1. Awad AI, Zawbaa HM, Mahmoud HA, Nabi EHHA, Fayed RH, Hassanien AE (2013) A robust cattle identification scheme using muzzle print images. In: Proceeding of the federated conference on computer science and information systems, pp 529–534
2. Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. In: European conference on computer vision, pp 404–417
3. Gaber T, Tharwat A, Hassanien AE, Snasel V (2016) Biometric cattle identification approach based on Weber's local descriptor and adaboost classifier. Comput Electron Agric 122:55–66
4. El Hadad HM, Mahmoud HA, Mousa FA (2015) Bovines muzzle classification based on machine learning techniques. Procedia Comput Sci 65:864–871
5. Kusakunniran W, Wiratsudakul A, Chuachan U, Kanchanapreechakorn S, Imaromkul T, Suksri-upatham N, Thongkanchorn K (2020) Biometric for cattle identification using muzzle patterns.

Int J Pattern Recognit Artif Intell 2056007

6. Lowe DG (1999) Object recognition from local scale-invariant features. In: Proceedings of the seventh IEEE international conference on computer vision, no 2, pp 1150–1157

7. Rublee E, Rabaut V, Konolige K, Bradski G (2011) ORB: an efficient alternative to SIFT or SURF. In: Proceedings of the IEEE international conference on computer vision, pp 2564–2571

8. Sian C, Jiye W, Ru Z, Lizhi Z (2020) Cattle identification using muzzle print images based on feature fusion. Proc Conf Ser: Mater Sci Eng 853(1):012051

9. Sultana MT, Dulhare UN, Rasool MS (2017) Feature point extraction by adaptive over-segmentation and feature point matching for effective digital image forgery detection. Int J Sci Eng Res 8(9)

10. Tharwat A, Gaber T, Hassanien AE (2015) Two biometric approaches for cattle identification based on features and classifiers fusion. Int J Image Min 1(4):342–365

11. Tharwat A, Gaber T, Hassanien AE, Hassanien HA, Tolba MF (2014) Cattle identification using muzzle print images based on texture features approach. In: Proceedings of the fifth international conference on innovations in bio-inspired computing and applications, pp 217–227

12. Tharwat A, Gaber T, Hassanien AE (2014) Cattle identification based on muzzle images using Gabor features and SVM classifier. In: Proceeding of conference on advanced machine learning technologies and applications, pp 236–247

A Secured Land Registration System Using Smart Contracts

M. Laxmaiah and B. Kumara Swamy

Abstract A secure land registration system using smart contracts is method to store land information which involves managing transactions of land titles, location, and ownership details. The blockchain technology is a distributed database ledger which stores all records of transactions available in the network. The important feature of the blockchain is establishing communication between trusted and untrusted third parties. On top of the blockchain, different distributed applications beyond cryptocurrencies can be deployed. One such application is a smart contract that executes codes between trusted parties. The secured land registration system is a powerful use case for blockchain technology which provides security for land documents. The system and method provide a solution to the problems related to the land registration and also develop a user friendly framework of smart contracts for the limitations of the people's knowledge about the blockchain and its applications.

Keywords Land records · Smart contracts · Blockchain technology

1 Introduction

The land is a valuable and immovable property. The estimation of the land value depends upon its area and with a developing population; its interest continues expanding, while its accessibility is getting restricted step by step. Authorization to land rights strongly affects occupations, mechanical, monetary, and social development. Landlords are an individuals having the superior in better access to own lands, ease to their own better access and for other monetary freedoms as a rights. Land proprietorship is comprehensively dictated by admittance to a land title, an archive

M. Laxmaiah (✉) · B. Kumara Swamy
Department of Computer Science and Engineering, CMR Engineering College, Medchal, Hyderabad, Telangana 50140, India
e-mail: datasciencehod@cmrec.ac.in

B. Kumara Swamy
e-mail: kumaraswamy.b@cmrec.ac.in

that states such possession. Having a sensible land title ensures the privileges of the titleholder against different cases made by any other person to the property.

Land proprietorship is resolved through different records, for example, enlisted sale deeds, local archives, and government review records. Land titles in India are indistinct because of different reasons, for example, inheritance issues from the Zamindari framework, flaws in the legal structure, and helpless organization of land records. Such mixed-up land records have prompted legal questions identified with land proprietorship and influenced the farming and land areas. Questioned land titles lead to an absence of straightforwardness in land exchanges making the housing market inefficient. Execution of new undertakings requires clearness on the proprietorship and estimation of land, the two of which become troublesome without clear land titles. Any foundation made ashore that is not without encumbrance can be possibly tested, later on, making its speculations dangerous.

Land possession in India is hypothetical. At present, the Transfer of Property Act-1882 gives that the privilege to animus capable property can be moved or sold exclusively by an enlisted report. Such records are enlisted under the Registration Act, 1908. Hence, in India, the mobilization of land or property refers to the enrollment of the exchange or sale deed, and not the land title. An enrolled agreement deed is not an administration assurance of land possession. This suggests that even bonafide property exchanges may not generally ensure proprietorship as a previous exchange of the property could be tested. During such exchanges, the onus of checking past proprietorship records so the far property is on the purchaser, and not the enlistment center. The land records can be kept up safely by utilizing blockchain innovation, keen agreements. Smart agreements are self-executing contracts regarding the arrangement among purchaser and dealer of the land being straightforwardly composed into lines of code.

The code and the arrangements contained in that exist across a dispersed, decentralized blockchain network. The smart agreements grant believed exchanges and arrangements to be completed among dissimilar, mysterious gatherings without the requirement for a focal power, overall set of laws, or outside authorization component.

They render exchanges recognizable, straightforward, and irreversible. Exchanges between parties in current frameworks are normally led in an incorporated structure, which requires the association of a confided outsider. Notwithstanding, this could bring about security issues and high exchange charges. Blockchain innovation has arisen to handle these issues by permitting untrusted elements to collaborate in a dispersed way without the inclusion of a confided outsider. Blockchain is a dispersed dataset that records all exchanges that have at any point happened in an organization. Blockchain was initially presented for Bitcoin, a shared computerized installment framework, yet then advanced to be utilized for building up a wide scope of decentralized applications. An engaging application that can be sent on top of blockchain is smart contracts. A smart contract is an executable code that sudden spikes in demand for the blockchain to encourage, execute, and implement the details of an understanding between untrusted parties. It tends to be considered as a framework that discharges computerized resources for all or a portion of the elaborate gatherings once the predefined rules have been met. Contrasted with customary agreements,

shrewd agreements do not depend on a confided outsider to work, bringing about low exchange costs.

Today, various kinds of blockchain strategies can be used to create strong agreements, yet Ethereum is the most well-known strategy. This is because Ethereum language underpins the Turing-fulfillment highlight which implies that it can answer any computational issue. Brilliant contracts have the accompanying highlights, for example, diminish the danger of enlisting inaccurate data, help with getting title deed and affirmation from the land vault of responsibility for the land, use remarkable advanced unique finger impression to manage and control the work process, the rightness of the record, and the request for the guidelines of approval. To address the above in this paper, we are building up a model for the protected land enrollment interaction to stay away from land enlistment-related issues.

The paper has 6 sections. This section of the paper depicts the introduction. Section 2 describes the related work. Section 3 clarifies the conventional arrangement of land registration. Section 4 describes the proposed land registration framework utilizing blockchain innovation. Section 5 brief up about the outcomes and discussion. Section 6 concludes the paper.

2 Literature Survey

In the literature survey, we noticed two things, smart contracts issues and smart contract-related topics. The smart contract-related issues are into different types such as coding, security, privacy, and performance issues of smart contracts.

2.1 Coding Issues in Smart Contracts

The first key issue in the smart contract is writing the correct smart contracts during the agreement. The smart contracts should be working as per the guidelines of the developers. Because they have valuable currency units. Some of the currency units may lose if the smart contracts are not working as per the direction of developers [1–4]. Termination of the smart contract is a second important issue in the blockchain. To handle this issue, Marino et al. have proposed a set of standards by not allowing the smart contracts to be terminated easily [3]. The third important issue is the lack of support to identify optimized smart contracts [5]. Smart contracts to store or compute require the cost for each operation. An under-optimized smart contract contains unnecessary expensive operations. Such operations cost may be higher, so that the user side, it becomes very expensive. To tackle this issue, Chen et al. [4] proposed seven programming patterns in smarts contracts which leads to unnecessary extra cost. They also proposed and developed a tool known as GASPER to optimize the cost of pattern execution in smart contracts. They examined Ethereum smart

contracts and found that most of the patterns are suffering from extra cost. The final issue is the complexity of the programming languages used for smart contracts [6].

2.2 Security Issues in Smart Contracts

In the literature survey, we noticed different security issues in smart contracts. These are transaction ordering, untrustworthiness in data feeds, criminal activities, dependency of timestamps, mishandling of exceptions. Atzei et al. [7] noticed several weaknesses in Etherium smart contracts. The important issue that occurs in transaction ordering dependency is that when two dependent transactions invoke the same contract may be included in one block.

The execution of the transaction order depends upon the miner. If the transactions are not executed in proper order, an adversary can launch attacks on the transaction [8, 9]. The second issue is mishandled exception vulnerabilities [10]. This problem occurs when a contract (caller) calls another contract (callee) without checking the value returned by the callee. When calling another contract, an exception (e.g., run out of gas) sometimes raised in the callee contract. This exception, however, might/might not be reported to the caller depending on the construction of the call function. Having not reported an exception might lead to threats as in the King of The Ether (KoET). The final issue is trustworthiness in data feeds; the smart contracts require guaranteed information providers from outside of the blockchain data source [8].

2.3 Privacy Issues in Smart Contracts

Privacy is the most important aspect in smart contract mechanism. In this, there are two types of privacy are available such as lack of transnational privacy and lack of data feeds privacy. In case of lack of transnational privacy, all the transactions and user balances are visible to all the public in the blockchain system [7, 9]. Due to this lack of privacy, many people could not consider financial transactions as secure. Kosba et al. have developed the Hawk tool which allows people to use it to maintain the privacy of the smart contracts without the implementation of cryptography [11]. Watanabe et al. have proposed a method, before deploying the smart contracts, they must be encrypted in the blockchain [8]. The people who are involved in smart contracts can access smart contact data with help of decryption keys. In the issue of data feeds privacy, the data which is required for smart contacts requests parties for feeds [10]. This request is known to everyone who is in the blockchain.

2.4 Performance Issues in Smart Contracts

In a blockchain system, smarts contracts are executed in sequential order. This will create a negative impact if the number of smart contacts is more to execute per second. Vukolic suggests that if the smart contacts are independent, they can be executed in parallel. By doing so, the performance of blockchain systems would be improved as extra contracts can be executed per second.

3 Traditional Land Registration System

Traditional and conventional land registration systems and allied methodologies are shown through various documents such as registered sale deeds, property tax receipts, and land survey documents conducted by government. Land ownership involving conventional documents and their maintenance have below given disadvantages are:

High Cost of Property Registration

During the registration time, the buyers have to pay the registration amount along with the stamp duty. Due to the high cost of registration, several properties are not registered properly. So the data which is shown in the records is mismatching with original data.

Poor Maintenance of Land Records

The land records contain a variety of data such as property title, location details, history of the transaction, mortgage information. This data is maintained in the form of documents in various departments at the village or district level. The departments are working independently with each other, and the data is not updated across all departments on time. So, many discrepancies are occurring in land records.

Land Titles Are not Clear

The land titles are not clear due to many reasons such as poor administration of land records, historical issues from the Zamindari system, and breaches in the legal framework. Such unclear land records are creating legal disputes among the people.

Effect of Poor Records on Future Property Transactions

It becomes difficult and big to access land records when data is spread across departments and has not been updated. One has to go back several years of documents, including manual records, to find any ownership claims on a part of the property.

This process may cause delays and it is inefficient. The traditional contract execution needs manual validation, to check the terms and conditions. Based on these conditions, we decide later with the use of steps according to the written agreement.

The process of traditional land registration is depicted in Fig. 1. The traditional land registration systems are facing different challenges such as time-consuming,

Fig. 1 Traditional land registration process
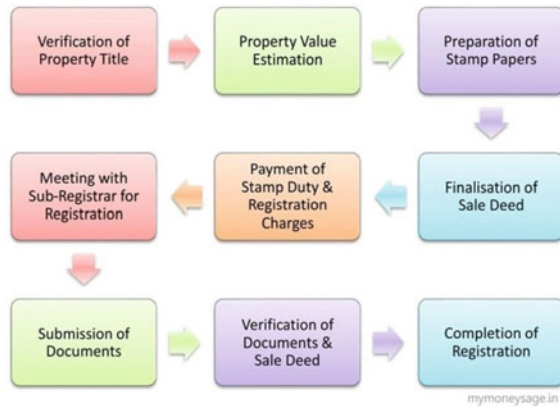
resource-consuming, costly, loss of privacy and integrity, a lot of scope for double-crossing, Communication gap, paper records, corruption, data history, vulnerability.

4 Our Approach to Land Registration System

The existing systems that employ blockchain, smart contracts, and other allied technology have certain lacunae such as coding, security, privacy, and performance issues to overcome the above disadvantages we proposed a new system which solves all the problems related to land registrations.

A typical process of smart contract is depicted in Fig. 2 and the system architecture is depicted in Fig. 3 which has three components such as private storage, account balance, and executable code. On the blockchain network, the state of the network is stored and when the contract is invoked, its state is updated automatically on the network. The contract code cannot be altered, once the contract is deployed into the blockchain. To run the smart contract, users are usually sending the transaction to the contract address. Then each transaction will be executed on every miner or consensus node in the network to reach consensus on its output. To run a contract, users can simply send a transaction to the contracts address. This transaction will then be executed by every consensus node (called miners) in the network to reach a consensus on its output. Subsequently, the state of contracts will be updated accordingly.

The contract can be based on the transactions that it receives, performs the read or write operations on the private storage media. It stores the money into the account

Fig. 2 Typical process of smart contracts

Fig. 3 System architecture
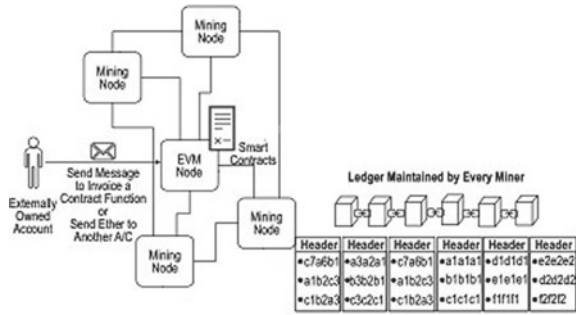
balance and sends or receives the money from other users. We can also create contracts for new users. The proposed system is to implement the concept of smart contracts and to record land dealings in the form of a distributed database ledger that stores records of land details in the blockchain. The proposed system is to implement the concept of smart contracts and to record land dealings in the form of a distributed database ledger that stores records of land details in the blockchain.

The block diagram of the smart contract of the land registration system is depicted in Fig. 4. Its components are briefed below:

Land registration system users: Buyers: are the people who search the system to purchase required land details. The buyers request to obtain the details of the sellers of the land and property titles, and the owners details. Sellers: are persons who use the system to sell the land property and manage land information, and transfer land titles to purchasers.
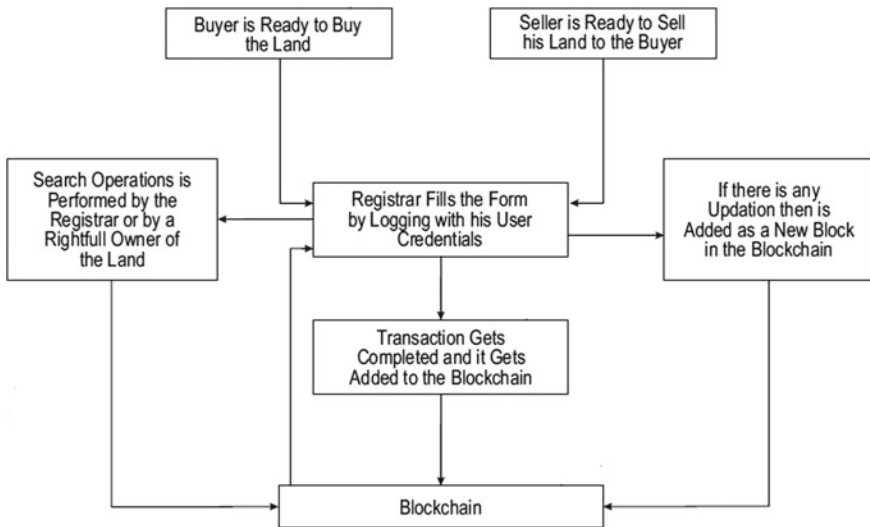
Fig. 4 Block diagram of land registration system

Registrar/Sub-Registrar: is a person, who will utilize the land registration framework to accomplish property requests, see land a document, confirm ownership details, and allocates land title ownership.

Registration of Users in the Land Registration Framework

The people who are willing to sell or buy the property can register in the land registering system. They will create their profile on the system by providing details such as name, address, designation, and id proofs. For every user, a hash code is generated to identify the user and store information in the blockchain network.

Uploading Details of Property by Sellers

The property sellers put their property details such as property images, documents, location of the land on the map, and the transactions related to the property are stored in the blockchain network. Whenever the details of the property are updated on the network, these details are made available to all the buyers on the platform.

Request to Access the Property by Buyers

The people who are willing to purchase the property will send a request to the sellers by sharing the specific details of the property with sellers. The property access request will be received from buyers about their interests. The sellers of the property can accept or reject the profile of the buyers. To ensure authenticity and traceability of the records, the transaction that is taking place between sellers and buyers are recorded on the blockchain network.

Registrar Will Get Information When the Sellers Approve the Request

On the network, whenever the landowner agrees to land ownership transfer, the registrar will receive a message from the seller for transferring land ownership to the buyer. The smart contract framework allows the registrar/land inspector to access the land documents for verification. Once the land documents verification is done, then the registrar will arrange a meeting between buyers and sellers to transfer the land ownership. The record of the meeting is added to the blockchain network to avoid disputes related to land in the future.

Registrar Will Transfer Land Ownership

The registrar is the final authority who verifies the documents submitted by both sellers and buyers and stores the land ownership details onto the blockchain smart contract land registration system. Buyers and sellers are going to sign on land transfer documents in front of the registrar online on the land registration platform. The documents which are signed by both parties are saved in a database, and their related transaction data is a record on the blockchain. The ownership transfer has occurred and the smart contracts system begins the fund transfer to sellers and the ownership transfer information to buyers of the property.

Validation and Authenticity Land Registration

Whenever any land disputes are taking place between the two parties, the authorized user of the system can upload their signed documents on the smart contract land registration system for verifying its authenticity and validated it on the system. Then, the hash code is compared between the times of signing the document with at the time of uploading the document on the system. If both hash codes are the same, the document is authenticated document and no modifications are allowed to the document.

5 Results and Discussion

The secure land registration system is a mobile-based application of smart contracts on the blockchain platform. The back end of the application of the smart contract has been implemented with the solidity framework. The front end of the application, the login page is developed using web page design technologies such as HTML and CSS. When the user accesses the application, the home page is opened. The home page is depicted in Fig. 5.

The users are interacting with the systems by entering the user name and password. The user login page has been developed using HTML and CSS. The client-side validations are carried out by using Java Scripts. Figure 6 depicts the login page where a user enters the username and password as two input to the system.

The land registration system takes the input from buyers and sellers in the form of land documents by containing details like owner name, land titles, location information and uploaded into the smart contract framework. The land registration form is shown in Fig. 7.

Fig. 5 Home page

Fig. 6 Login page
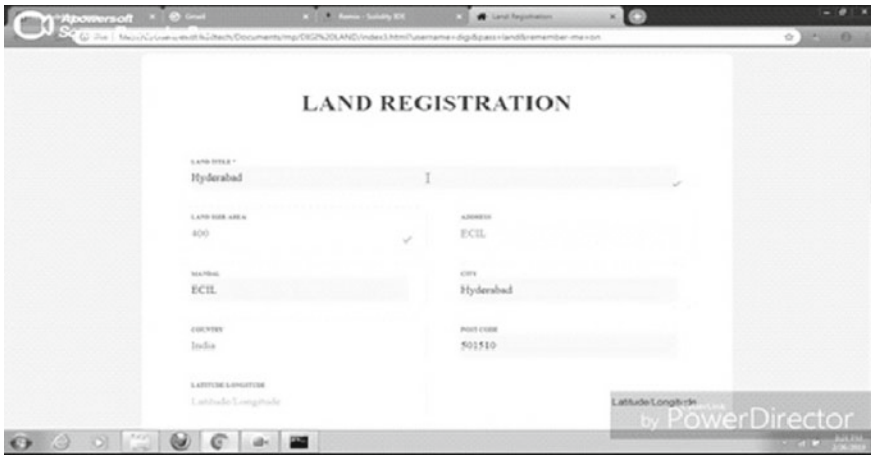
Fig. 7 Land details registration form

The process of every transaction entered into the blockchain network is to provide security for digital documents. The transaction information of land registration documents is uploaded into blockchain technology. During transaction time, it generates hash code for it. The hash code helps to avoid tampering with the land documents if anyone is accessing them. The registration of transactions is stored in a blockchain is shown in Fig. 8.

6   Conclusion

In this paper, we developed a system to provide value by improving processes associated with land registration and real estate transactions. The system eliminates the need for physical archives of contracts and files, redundancy of the transaction data in the land registry, and the mortgage deed registry. It provides greater security to the

Fig. 8 Entry of the transaction into blockchain and acknowledgment

users of the system, in part because validation of the purchasing contracts and ownership can be done independently, faster and more transparent transactions, increased liquidity of real estate since it can be sold soon by the owner, making it possible to receive an automatic confirmation of final land title at the date of transaction. The system also eliminates the possibility of selling a property more than once and making it very difficult to steal property.

References

1. Lewis A. A gentle introduction to smart contracts. Available at: https://bitsonblocks.net/2016/02/01/agentle-introduction-to-smart-contracts/. Accessed 25 Feb 2017
2. Frantz CK, Nowostawski M (2016) From institutions to code: towards automated generation of smart contracts. In: 2016 IEEE 1st international workshops on foundations and applications of self* systems (FAS*W). IEEE, pp 210–215
3. Marino B, Juels A (2016) Setting standards for altering and undoing smart contracts. In: International symposium on rules and rule markup languages for the semantic web. Springer, pp 151–166
4. Chen T, Li X, Luo X, Zhang X (2017) Under-optimized smart contracts devour your money. In: 2017 IEEE 24th international conference on software analysis, evolution, and reengineering (SANER). IEEE, pp 442–446
5. Natoli C, Gramoli V (2016) The blockchain anomaly. In: 15th international symposium on network computing and applications (NCA). IEEE, pp 310–317
6. Luu L, Chu D-H, Olickel H, Saxena P, Hobor A (2016) Making smart contracts smarter. In: Proceedings of the 2016 ACM SIGSAC conference on computer and communications security, CCS '16. ACM, pp 254–269
7. Atzei N, Bartoletti M, Cimoli T (2017) A survey of attacks on Ethereum smart contracts (SOK). In: International conference on principles of security and trust. Springer, pp 164–186
8. Watanabe H, Fujimura S, Nakadaira A, Miyazaki Y, Akutsu A, Kishigami JJ (2015) Blockchain contract: a complete consensus using blockchain. In: 2015 IEEE 4th global conference on consumer electronics (GCCE). IEEE, pp 577–578
9. Vukoli M (2017) Rethinking permission blockchains. In: Proceedings of the ACM workshop on blockchain, crypto currencies, and contracts, BCC '17. ACM, pp 3–7
10. Zhang F, Cecchetti E, Croman K, Juels A, Shi E (2016) Town crier: an authenticated data feed for smart contracts. In: Proceedings of the 2016 ACMSIGSAC conference on computer and communications security, CCS '16. ACM, pp 270–282
11. Kosba A, Miller A, Shi E, Wen Z, Papamanthou C (2016) Hawk: the blockchain model of cryptography and privacy-preserving smart contracts. In: 2016 IEEE symposium on security and privacy (SP). IEEE, pp 839–858

A Constructive Feature Grouping Approach for Analyzing the Feature Dominance to Predict Cardiovascular Disease

K. S. Kannan, A. Lakshmi Bhargav, A. Anil Kumar Reddy, and Ravi Kumar Chandu

Abstract Predicting the resistivity of an individual is essential for the optimal and prompt treatment against cardiovascular disease (CVD) in the earlier stage, which recommends the requirement for productive risk evaluation tools. The data-driven-based approach can predict every individual's risk by handling the crucial data patterns. To facilitate clinically applicable CVD prediction resolving the missing data patterns and interpretability issues using machine learning (ML) approaches. Here, a multi-tier model is proposed for mining missing data patterns. Initially, data fusion is adapted to describe the block-wise data patterns. It enables patient data (1) grouping-based feature learning and imputation of missing data and (2) prediction model considering the data availability. The feature selection process uses group characterization to uncover the risk factors. Then, the boosting model is generalized for identifying the patient's sub-group. The experimentation is done on an online available UCI ML dataset to demonstrate the significance of the model compared to various other approaches. The model attains 99% prediction accuracy, which is substantially higher than other approaches.

Keywords Cardiovascular disease · Risk evaluation · Missing data patterns · Grouping · Data imputation

1 Introduction

Cardiovascular illnesses seem to be the most existing disorders globally, with the greatest fatality rate. They have grown quite frequently throughout time, and they are now overburdening the medical systems of societies. Hypertension, family medical history, depression, age, sexuality, cholesterol, BMI, and obesity rates are key risk factors for heart disease attacks [1]. Experts have investigated numerous ways for

K. S. Kannan · A. Lakshmi Bhargav (✉) · A. Anil Kumar Reddy
Department of CSE, CMR Engineering College, Hyderabad, India
e-mail: albhargav543@gmail.com

R. K. Chandu
Department of AI/ML, CMR Engineering College, Hyderabad, India

early detection based on these criteria. However, given the inherent structural failure and the dangers of high blood pressure, the offered procedures' effectiveness needs some refinement [2]. The approach addresses incorrect values (using the mean skills that will help) and informational unbalance initially (via Synthetic Minority Oversampling Technique—SMOTE). The variable importance approach is used to choose features [3]. Finally, an assembly of regression model and K-Nearest Neighbor (k-NN) classifications is presented for a more accurate assessment. Standard benchmark databases (Framingham, Heart Failure, and Cleveland) are used to validate the methodology, with accuracies of 99.1, 98.0, and 95.5% attained in each case [4]. Finally, a comparison of MaLCaDD projections to current framework techniques shows that MaLCaDD predictions are more accurate (with a smaller collection of characteristics) than existing framework approaches [5]. As a result, MaLCaDD is very trustworthy and may detect cardiac disorders early in the actual world.

The current era's heavy workload leads to an unsustainable diet, which leads to anxiety and sadness. To cope with these circumstances, people often turn to cigarettes, alcohol, and using substances in the extreme [6]. Many severe illnesses, such as cardiac disorders and conditions, are caused by these factors. CVD is a leading cause of mortality based on the WHO report [7]. CVDs are accountable for over 30% of all fatalities worldwide. Early detection of these disorders is critical to perform preventative steps before anything more catastrophic occurs. The phrase "cardiovascular disease" (CVD) refers to a disorder that involves the cardiac or blood arteries. Cardiovascular disease, stroke/transient ischemic attack (TIA/MiniStroke), circulatory diseases, and valvular illnesses are the four primary kinds of CVDs [8]. The specific origin of CVDs is undetermined; however, several risk factors, such as high cholesterol, nicotine, diabetes, body mass index (BMI), lipids, age, and family medical history, are linked to cardiovascular diseases. Different people are affected by various circumstances. CVDs are caused by various variables, including age, genetics, pressure, and an addictive personality [9]. The main problem is accurately forecasting these conditions promptly to minimize death rates by using the appropriate treatment and other interventions.

Researchers have developed numerous techniques for the prediction of CVDs throughout time. On the suggestion of the Spanish Society of Cardiology, the SCORE risk chart and REGICOR risk score are used in Spain for cardiovascular risk reduction. Similarly, researchers from the University evaluated five approaches for determining morbidity in patients with coronary artery disease (CHF). To predict CVDs, prospective coronary Münster (PROCAM) research is established. The author in [9] utilized ECG signals as one of the popular reasons for predicting cardiovascular illness. Cardiovascular imaging could also provide image-based diagnoses and forecast CVDs [9]. Other factors have also been employed by certain studies to forecast vascular disorders.

Using various information and approaches, scholars have reported multiple algorithms to predict CVDs throughout the last generation. Heart disease, Cleveland, Framingham, and Cardiac Disease are some of the most often used statistics for CVD forecast. These databases comprise several characteristics used to diagnose CVDs. Modifiable and non-modifiable risk factors are among the variables that contribute

to coronary heart disease. Gender, race, and personal history are present in non-characteristics [10]. Lifestyle changes, such as tobacco, leading to an unhealthy diet, hypertension, and high cholesterol, may be modified and reduced with the right measures and medicine. Much information has been constructed by considering certain features, and academics have put a lot of work into some of these databases. The Framingham collection [10], collated against all of these parameters, is one of the most well-known databases. Several academics have utilized this data collection to evaluate their forecasting model.

Different machine learning [11] based strategies for measuring the prognosis of CVDs were designed in the provided study environment. On the other hand, specialists are primarily concerned with feature extraction methods and SVM classifiers, neglecting the issue of misclassification. Cultural imbalances have a significant impact on the segmentation application's performance. Secondly, when the material is not harmonized, many characteristics are necessary for predicting. It considerably increases the computing complexity of the solution, rendering it unsuitable for use in a real-world setting. Furthermore, conventional feature extraction techniques are improved to minimize the skill required while maintaining the required correctness. Similarly, pre-existing classifying prediction accuracy is enhanced to produce trustworthy findings. To summarize, an integrative machine learning paradigm for coronary illnesses is urgently needed, with data balance, effective feature identification, and enhanced classifier carried out uniformly. It improves coronary heart disease forecasts and minimizes the curse of dimensionality [10]. This work concentrates on modeling an efficient data pattern analysis model to improve the quality of prediction accuracy.

The work is structured as: Sect. 2 provides a detailed analysis of various prevailing approaches. The methodology is explained in Sect. 3, with the numerical outcomes in Sect. 4. The research summary is given in Sect. 5.

2 Related Works

It is a significant study issue in the current medical sector to measure future illness risks and provide proper diagnostic forecasts for individuals. Potential diagnostic, the chance of particular illness, mortality rate, hospital readmission, acute techniques, and survival time are core priorities of clinical predicting jobs. The major focus of this study is on cardiac risk prognosis. Most of the strategies for predicting blood cholesterol levels in previous studies are focused on linked cohort studies. Everett et al. [11] studied the connection between the diagnostic NT-proBNP (N-terminal pro-B-type natriuretic peptide) and CVDs in population analysis of 1821 CVD reported cases. Moreover, it is shown that NT-proBNP might increase CVD risk prediction ability. Author et al. [12] introduced a unique technique for extracting features from EHRs built on relation regularization, which considerably increased the efficiency of estimation for myocardial ischemia (ACS). EHRs are used to construct a reliable and accurate risk prediction system for readmission rates. However, because

HIS EHRs are often high-dimensional and diverse, input image becomes difficult. Human cultural methodologies are routinely used to develop feature maps in many extant estimation techniques. Pike et al. [26] retrieved associated features from EHRs based on specialized cardiovascular disease risk scores, then analyzed the completion of different prediction models, such as the QRISK II Score and the Framingham Risk Score (FRS). In their prediction model is used standard risk indicators from the FRS. In addition, they used the statistic method to analyze specific variables from EHRs to enhance cardiovascular risk projection. Furthermore, traditional feature development relies heavily on specialist medical knowledge. Some new techniques [13] have now been presented to implement autonomous feature extraction to lessen the need for manual intervention. Machine learning algorithms are frequently incapable of comprehending sequence EHR data because a database schema is often rich in video sequences.

Most studies have boosted accuracy through image segmentation and classical SVM classifiers. Missing values in the database, which have a significant impact on the model's correctness (since missing values lower the number of observations in the statistics, resulting in an unsuccessful model), have received very little attention [14]. Furthermore, class imbalance (in which samples from one class are significantly smaller than sampling from the other class/classes) has not been well addressed in earlier studies intended to improve correctness. As a result, difficulties such as missing data and class unbalance must be addressed well before the classification method is proposed. Similarly, database characteristics significantly impact the final and computing efficiency of the training phase. As a result, while doing feature extraction, the most important aspect of any training process is to choose the appropriate subset of attributes. Even though research teams have proposed two feature selection methodologies and SVM classifiers for the available dataset, the pre-processing stage requires marked enhancement to select the appropriate feature subsets that make a significant contribution to accurate determination with greater accuracy [15]. To decrease the powerful cryptographic of suggested classifications, the subset of features approach must extract the absolute necessities of important features that can aid in producing valid predictions. Furthermore, the categorization method requires improvements in reliability.

3 Methodology

In this research article, the evolutionary classifier model is a significant cause for heart disease prediction and constructs a model with maximal possible accuracy. The dataset is attained from the UCI machine learning repository [15]. It comattributes real-time samples with 14 diverse attributes where '1' class and '13' predictors such as electrocardiogram results, chest pain types, and blood pressure (see Table 1). Here, pre-processing the UCI heart disease dataset is done with data normalization.

Table 1 Dataset attributes and descriptions

S. No.	Attributes	Specification	Descriptions
1	Age	Age	Age (years)
2	Sex	Sex	1 → male; 0 → female
3	Chest pain	c_p	1 → typical angina 2 → atypical angina 3 → non-angina 4 → asymptomatic
4	Rest BP	Trestbps	Systolic blood pressure during resting (mm Hg)
5	Serum cholesterol	Chol	Serum cholesterol (mg/dl)
6	Fasting blood sugar	FBS	Blood sugar during fasting > 120 mg/dl 0 → false 1 → true
7	Rest ECG	Respect	0 → normal 1 → ST-T wave abnormality 2 → left ventricular hypertrophy
8	Maximal heart rate	Thatch	Attainment of maximal heart rate
9	Exercise-induced angina	Exchange	Induced angina (exercise) 0 → no 1 → yes
10	ST depression	Old peak	Depression induced due to relative exercise to rest
11	Slope	Slope	Peak exercise slope 1 → upsloping 2 → flat 3 → downsloping
12	No. of vessels	C_a	No. of significant vessels (0–3)
13	Thalassemia	Thal	Defect types 3—normal 6—fixed defect 7—reversible defect
14	Number of class attributes	Class	Predicting heart disease status 0—nil 1—low risk 2—potential risk 3—high risk 4—extremely high risk

Pre-processing

It is the most general approach used for performing pre-processing. It converts the image values into a specific range, i.e., 0 and 1. The normalization approaches considered here are z-score normalization and zero means. It is expressed as in Eq. (1):

$$X'_i = \frac{X_i - \text{mean}\,(X)}{\text{SD}(X)} \tag{1}$$

Here, X'_i is normalized data; mean(X) specifies mean value, and SD (X) specifies standard deviation of input X. The SD is mathematically expressed as in Eq. (2):

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(X_i - \text{mean}\,(X))^2} \tag{2}$$

Here, σ specifies the standard deviation for the provided input values.

Data Pattern Analysis

The clinical features are organized into clusters related to various clinical evaluation measures based on the UCI ML dataset. The medical experts provide the available diagnostic clinical records with inherent similarities among the individuals' conditions. The feature learning process is well-suited for fusing the clinical features and makes the incomplete and complex data easily be handled and interpreted. The feature matrix $n * d$ is represented using the d—dimensionality of the n patients acquired from the clinical UCI dataset to categorize the missing data patterns. The feature matrix M_{ij} entries are set as 1 or 0 for missing data or not. The constant values of fused data patterns are predicted through the fusing process. The missing and observed features are labeled with distinct variations. Based on the clinical feature availability, the indicator matrix of individuals with the UCI dataset can organize patients and features simultaneously into clusters. This research works sequentially in integrating the clustering process. The clustering process is enabled based on the grouping of feature selection with a certain prediction model. Assume that the feature groups (1 and 2) are chosen for patients' sub-group 1, and feature groups (2 and 3) are chosen for patients sub-group 2. The prediction models are trained for patients sub-group based on the clinical data availability (see Fig. 1).

Feature selection plays a substantial role in enhancing computational efficiency, learning performance and avoiding over-fitting. With the availability of the feature groups related to various clinical evaluation measures, this work intends to choose or not to choose the features with the same group concurrently. Some generally used approaches like filtering methods are based on the knock-off and mutual information for feature selection to avoid grouping feature structures. On the contrary, the grouping is formed based on the k—disjoint groups. The model parameters are regularized based on the group coefficients. Here, a localized prediction model is trained for every patient sub-group to enhance the prediction accuracy and benefit individual group structures. The proposed feature learning model is intelligent and represents higher accuracy than various other approaches. The enhancements of the anticipated model over conventional approaches include the integration of learning approaches and the automatic prediction of pair-wise feature interactions as in Eq. (3):

$$g(E[y]) = \beta_0 + \sum f_j(x_j) + \sum f_{ij}(x_i, x_j) \tag{3}$$
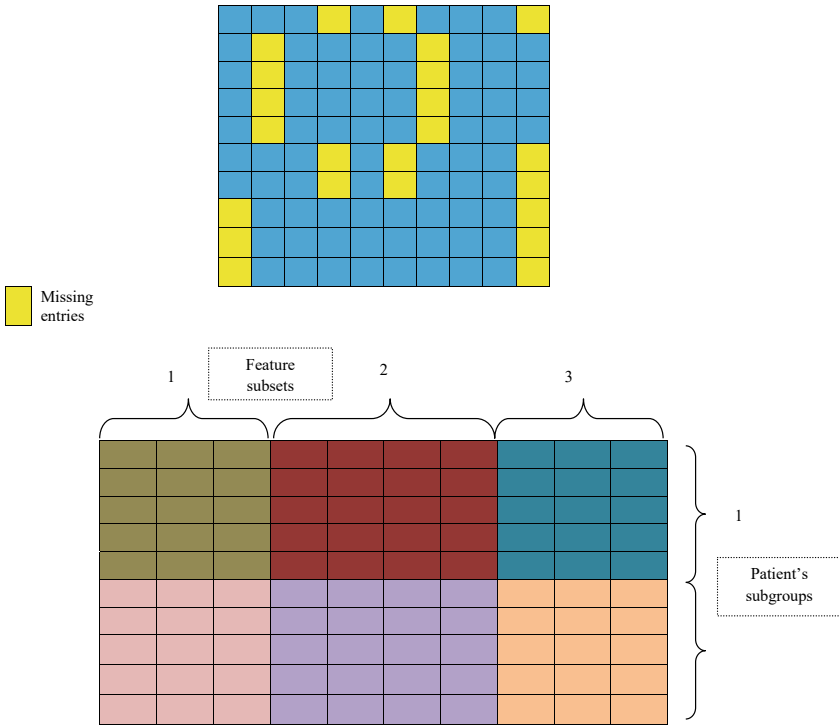
Fig. 1 Data pattern analysis

This model is effectively learned using gradient boosting with a tree-like ensemble model.

4 Numerical Results and Discussion

This work considers six performance metrics to compute the performance of the anticipated model. The potential model output includes True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP). Here, TP and TN are depicted as the number of samples appropriately classified as 'positive' (occurrence of heart disease), and negative specifies the absences of heart disease, respectively. Then, FP and FN specifies the number of samples inappropriately classified as 'positive' (heart disease occurrence) when it is usually 'negative' (absence of heart disease) and inappropriately classified as 'negative' (absence of heart disease) when it is actually 'positive' (heart disease occurrence), respectively. The execution of the proposed classifier model is done in MATLAB 2018a using various learning libraries. The experimentation is done in Intel-core i7 process, 16 GB RAM on Windows 10 OS. Additionally, the performance metrics are evaluated, and accuracy is expressed as in

Eq. (4):

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \qquad (4)$$

Precision is expressed as in Eq. (5):

$$Precision = \frac{TP}{TP + FP} \qquad (5)$$

Sensitivity/recall/True Positive Rate (TPR) and F1-score is expressed as in Eqs. (6) and (7):

$$Recall = \frac{TP}{TP + TN} \qquad (6)$$

$$f = \frac{2pr}{p + r} \qquad (7)$$

Table 2 depicts the comparison of overall performance metrics. The comparison is made among the existing LR, RF, L-GBM, SVM, BRT, DFN, and the proposed (see Figs. 2, 3 and 4). The metrics like accuracy, sensitivity, specificity, F1-score, time, and AUROC are measured and compared with these methods. The accuracy of the anticipated model is 99% which is 9.44, 18.41, 2.55, 2.82, 12.35, and 0.35% higher than other approaches during data pattern analysis. The sensitivity of the anticipated model is 96% which is 26, 27, 30, 39, and 1% higher than other approaches. The specificity of the anticipated model is 97% which is 29, 1, 26, 25, 29, and 3% higher than other approaches. The F1-score of the anticipated model is 45% which is 10% higher than LR, RF, L-GBM, SVM, and BRT model and 5% higher than the DFN model. The execution time is 0.03 min which is higher than other approaches. The AUROC is 97% higher than other approaches.

5 Summary

This work investigates data-driven approaches for CVD prediction, and the proposed multi-tier approach integrates the grouping and feature learning process. The proposed interpretable model measures the data's incompleteness with the enhanced prediction process. The model shows that grouping or clustering can exploit the missing data patterns by aggregating the features and patterns simultaneously. The supervised learning approach is proposed to improve the underlying nature of real-time data. Even with the improved performance of the prediction model, there are diverse limitations that need to be resolved in the future. There is a need for sufficient data to ensure the efficacy of the prediction model. The optimization models and the classification approach can enhance the model performance.

Table 2 Overall comparison of the performance metrics

Dataset	ML algorithm	Accuracy (%)		Sensitivity	Specificity	F1-score	No. of features		Time (min)	AUROC
		Without feature	With feature				Without feature	With feature		
UCI machine learning repository	LR	88.56	95	0.70	0.68	0.35	14	12	1.65	88.56
	RF	79.59	92	0.69	0.96	0.35	14	9	0.33	79.96
	L-GBM	95.45	94	0.66	0.71	0.35	14	12	4.05	95.45
	SVM	95.18	94	0.68	0.72	0.35	14	12	37.98	95.18
	BRT	85.65	93	0.69	0.68	0.35	14	10	2.58	85.65
	PDF	97.65	98	0.95	0.94	0.40	14	7	0.04	96.85
	Proposed	**98**	**99**	**0.96**	**0.97**	**0.45**	**14**	**6**	**0.03**	**97**
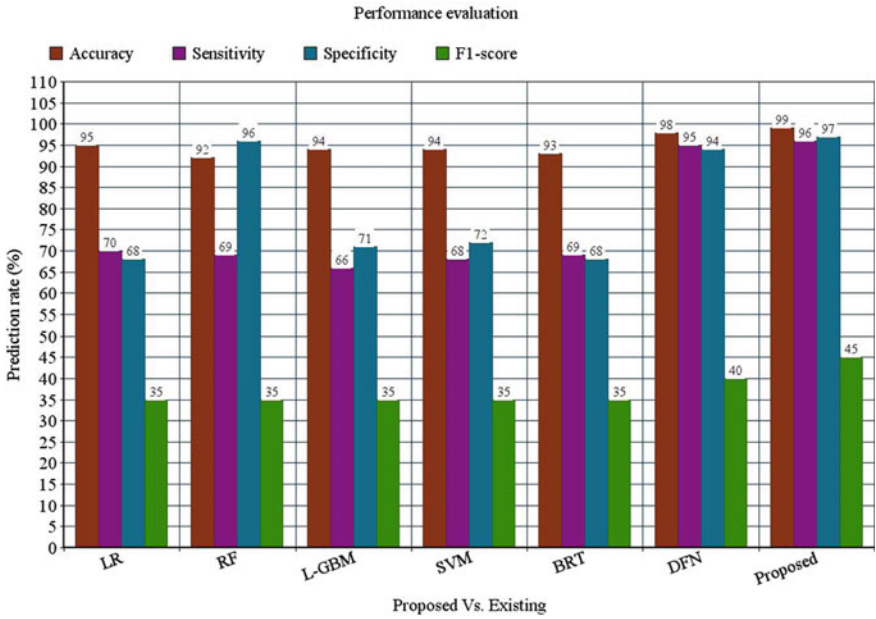
Fig. 2 Performance metrics evaluation
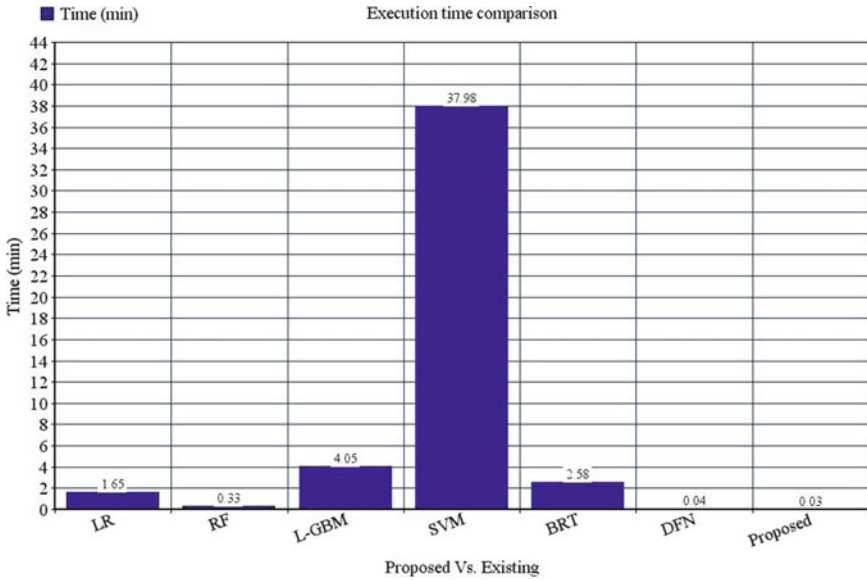
Fig. 3 Execution time comparison
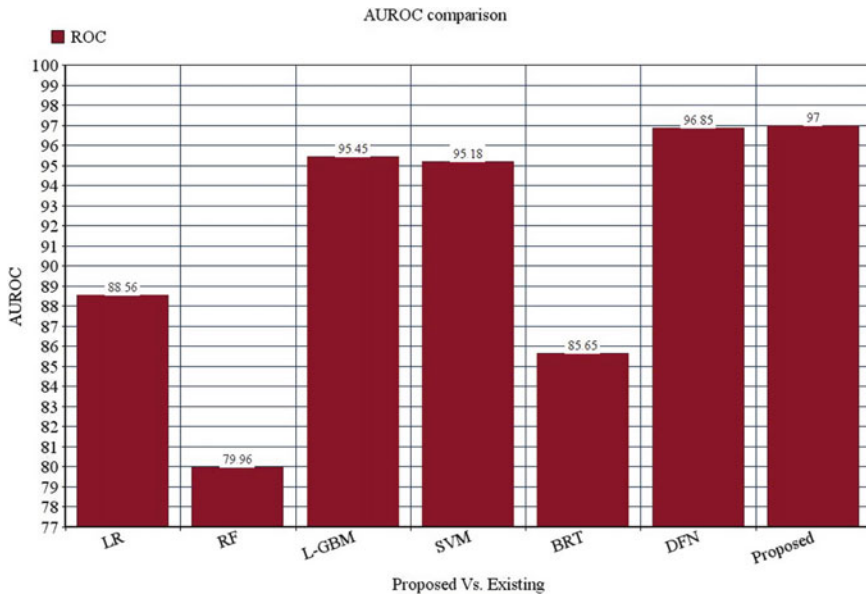
Fig. 4 AUROC comparison

References

1. Jensen LJ, Brunak S (2012) Mining electronic health records: Towards better research applications and clinical care. Nat Rev Genet 13(6):395–405
2. Li Y, Bai C, Reddy CK (2016) A distributed ensemble approach for mining healthcare data under privacy constraints. Inf Sci 330:245–259
3. Khedr AM, Aghbari ZA, Kamel I (2018) Privacy-preserving decomposable mining association rules on distributed data. Int J Eng Technol 7(3–13):157–162
4. Khedr, Bhatnagar R (2014) New algorithm for clustering distributed data using K-means. Comput Inf 33:1001–1022
5. Subhashini, Jeyakumar MK (2017) OF-KNN technique: an approach for chronic kidney disease prediction. Int J Pure Appl Math 116(24):331–348
6. Nathan, Kumar PM, Panchatcharam P, Manogaran G, Varadharajan R (2018) A novel Gini index decision tree data mining method with neural network classifiers for heart disease prediction. Des Autom Embedded Syst 22(3):225
7. Hsu, Manogaran G, Panchatcharam P, Vivekanandan S (2018) A new approach for prediction of lung carcinoma using backpropagation neural network with decision tree classifiers. In: Proceedings of the IEEE 8th international symposium on cloud and services computing (SC), Nov 2018, pp 111–115
8. Ramasamy, Nirmala K (2020) Disease prediction in data mining using association rule mining and keyword-based clustering algorithms. Int J Comput Appl 42(1):1–8
9. Bakar, Kefli Z, Abdullah S, Sahani M (2011) Predictive models for dengue outbreak using multiple rule-based classifiers. In: Proceedings of the international conference on electrical engineering and informatics, 2011, pp 1–6
10. Hariharan, Umadevi R, Stephen T, Pradeep S (2017) Burden of diabetes and hypertension among people attending health camps in an urban area of Kancheepuram district. Int J Community Med Public Health 5(1):140

11. Tun, Arunagirinathan G, Munshi SK, Pappachan JM (2017) Diabetes mellitus and stroke: a clinical update. World J Diabetes 8(6):235–248
12. Ley, Hamdy O, Mohan V, Hu FB (2014) Prevention and management of type 2 diabetes: dietary components and nutritional strategies. Lancet 383(9933):1999–2007
13. Meng, Huang Y-X, Rao D-P, Zhang Q, Liu Q (2013) Comparison of three data mining models for predicting diabetes or prediabetes by risk factors. Kaohsiung J Med Sci 29(2):93–99
14. Bashir, Qamar U, Khan FH (2016) IntelliHealth: a medical decision support application using a novel weighted multi-layer classifier ensemble framework. J Biomed Inform 59:185–200
15. Nai-Arun, Moungmai R (2015) Comparison of classifiers for the risk of diabetes prediction. Procedia Comput Sci 69:132–142

High-Band Compact Microstrip Patch Antenna for 5G Wireless Technologies

Aditya Prajapati and Sweta Tripathi

Abstract We live in a scientific and technological age, and technologies are growing at a rapid pace. Microstrip antennas are one of the well-known antennas and are widely used because of its diminutive size and can be effortlessly put into practice on printed circuit board. This paper presents a high-band microstrip antenna of dimension $(5 \times 5 \times 1)$ mm^3 for 5G wireless mobile communications ranging up to 40–45 GHz. The computations of the antenna and its substrate, electrical and physical parameters are calculated by using microstrip line and patch calculator. The antenna frameworks are optimized and simulated by using Ansys HFSS software. The planned antenna has the advantages of simple fabrication and compactness which can be used in various wireless appliances, future IoT technologies and 5G wireless systems.

Keywords Microstrip antenna · Mobile communication · Wireless communication · 5G

1 Introduction

Nowadays, technologies are advancing at an incredibly rapid rate. In modern 5G wireless communication systems, microstrip antenna plays a vital role. Microstrip antenna can be easily implemented on portable devices. Typically, the transmitter or receiver of the antenna is linked by a very thin ductile metal sheet transmission line. From last ten years, microstrip antennas have become famous in modern communication due to their emaciated planar shape which allows them to be integrated into the surfaces of user merchandise, airliner and weaponry. Their simplicity of production using printed circuit techniques and the ability to add strenuous plans such as incorporation of microwave integrated circuits to the antenna itself to create active antennas, presents some of their fascinating properties [1].

A. Prajapati (✉) · S. Tripathi
Department of Electronics and Communication Engineering, Kanpur Institute of Technology, Kanpur, India
e-mail: adityaprajapati648@gmail.com

Antenna's world is very exciting and the microstrip patch antenna is very famous because of its characteristics and advantages. It can be in different shapes, compact, broadband, multiband, dual polarization, circular polarization, linear and planar arrays (series and parallel feeds). Advantages for which microstrip antenna is famous are less weight, less volume, flexible printed circuit boards are available which can be wrapped around a missile, simple configuration, it can be put underneath the aeroplanes and it can be put inside a pager or mobile phone. In Refs. [2–12], multiple researches are done on microstrip patch antenna on different applications that can help in defence systems and also for commercial purposes.

The 5G spectrum consists of radio frequencies in the sub-6 GHz range in addition to millimetre-wave (mmWave) frequencies of 24 GHz and above. There is couple of frequency bands for 5G networks. The long term evolution (LTE) frequency range is included in frequency range which runs from 448 MHz to 6 GHz is acknowledged as the sub-6 GHz. Another frequency range spans from 24 to 52 GHz which is acknowledged as the mmWave spectrum [13–16].

There are different shapes of antennas which have been extensively studied previously. Multiple patch shapes of antennas are discussed in Refs. [17–21]. Any different shape of the antenna will give different bandwidth. Suppose U-shape, C-shape and E-shape patch antennas and many more, so there is only difference which is the distribution of current for different frequency available, to get better bandwidth. But it makes the manufacturing quite complex. So in this manuscript, the planned antenna is in uncomplicated rectangular shape.

In this paper, we are enhancing the bandwidth and efficiency of the antenna and reducing the size of the antenna, comparing with the studies mentioned in Refs. [22–25]. If the antenna is loaded with slots in the ground plane, then the antenna is proficient to obtain a broad bandwidth. This is because it lessens the return loss of the antenna. For that reason, a cut is made on the patch and the microstrip inset feed method is used to feed the planned antenna while considering the input impedance equals to 100 Ω. The result of planned antenna shows the 86.61% efficiency with maximum gain of 3.90 dB which is approximately 4 dB in Z-axis direction, and the bandwidth obtained is equal to 4.64 GHz.

2 Literature Review

This section focuses on some recent advances on microstrip antenna for 5G wireless communication application.

Chauhan et al. [22] discussed the mmWave antenna for 5G mobile wireless systems. The antenna comprises of 1×1 arrays of two rectangular radiating elements. A single layer 'RT/Duroid 5880' substrate is used. It can operate on 28 and 38 GHz frequency. At resonant frequency 38.11 GHz, the input reflection coefficient for the antenna element is − 42.78 dB. The antenna's results are simulated in ADS software. The electric field radiation pattern of the array in ADS software yields a high gain of 9.025 dB, and the directivity obtained is 10.0336 dB. In modelling of the $S(1, 1)$

parameter for the frequency band 37–38 GHz, a maximum − 10 dB bandwidth of 1.27 GHz was attained. Efficiency of the antenna reported is 83%.

Güneşer and Şeker [23] discussed the microstrip antenna model which is compact in nature for 5G wireless communication. When examined using the return loss parameter, the suggested antenna exhibits one resonance frequency in the frequency range of 23–33 GHz. The return loss value is − 20.26 dB at the required centre frequency of 28 GHz. The bandwidth is determined to be 656 MHz among the two sites with return loss values of − 10 dB around the centre frequency. The suggested antenna resonates between 27.64 and 28.30 GHz, which corresponds to millimetre communication. However, efficiency is not discussed in their study.

Kapoor et al. [24] explain the wideband-printed antenna for 5G applications. The developed antenna is intended for use in C-band of the super high frequency spectrum, as well as from 3.3 to 4.2 GHz band of frequency range, and 5G bands such as sub-6 GHz. The maximum gain obtained is 2.5 dB at 3.83 GHz frequency. The VSWR is 1.70 dB at 3.83 GHz. This antenna can support sub-6 GHz band's wireless appliances. The simulation is done with the help of HFSS software. However, efficiency and directivity of this antenna is not discussed.

Przesmycki et al. [25] explain the broadband microstrip antenna which is able to be operating at 28 GHz. The antenna has a small footprint, measuring (6.2 × 8.4 × 1.57) mm³. The 'RT Duroid 5880' material with a dielectric coefficient of 2.2 and thickness of 1.57 mm was employed as a substrate for the antenna construction. The antenna has a low reflection coefficient of 22.51 dB, a high energy gain value of 3.6 dB, excellent energy efficiency of around 80% and wide working range of 5.57 GHz. However, the directivity of the antenna is not discussed in this reference. In rejoinder to the increased need for mobile data and devices, a rectangular microstrip antenna has been developed for 5G accomplishments. This antenna has a resonance frequency of 28 GHz and a reflectivity of − 22.50 dB.

3 Antenna Configuration and Design

In this segment, the planned antenna's configuration and design are discussed. The antenna consists of a ground sheet, FR4 substrate and rectangular patch. The dimensions of the antenna are explained below:

Figure 1 depicts the desired dimension of the planned antenna. The antenna dimension is (5 × 5 × 1) mm³ where $L = 5$ mm, $W = 5$ mm, $h = 1$ mm and the patch size is (1.04 × 2.03) mm² where $l = 1.04$ mm and $w = 2.03$ mm. The patch is printed on a FR4 substrate having height $h = 1$ mm, relative permeability $\mu_r = 1$, dielectric loss tangent $= 0.02$, relative permittivity $\varepsilon_r = 4.4$ and mass density $= 1900$. The antenna is fed by microstrip line inset feed technique along with input impedance, i.e. $Z_{in} = 100\ \Omega$. The planned antenna is designed in Ansys HFSS software and successfully simulated in it. The FR4 substrate is mounted over the ground sheet, and thickness of the ground sheet is 0.035 mm. The antenna patch and ground are assigned as a perfect electric. All the dimensions are revealed in Fig. 1a and b clearly.
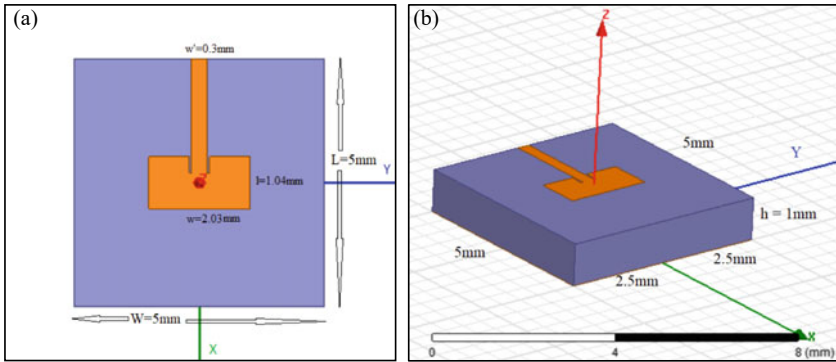
Fig. 1 a Top view and **b** side view of the planned antenna

It is an uncomplicated design so that it can be easily manufactured. The size of the planned antenna is very tiny, and it can be fitted in any small portable device. The resonance frequency is 45 GHz. With reference of resonance frequency 45 GHz, the measurements of the planned antenna are taken. The measurements of rectangular patch are calculated in 'em:talk microstrip patch calculator' and the measurements of feed line of the antenna are calculated in 'em:talk microstrip line calculator'.

This section focuses on some recent advances on microstrip antenna for 5G wireless communication application.

3.1 Calculation of Cut Size

$$\textbf{Formula}: \quad Z_{\text{in}}(y = y_0) = 243 \cos^2(\pi y_0/l)\text{Radian}$$

where Z_{in} is input impedance, l is the length of the patch and y_0 is the length of the cut size.

Length of the feed line will be available space on the antenna. With the help of this formula, we have to find only width size for feeding the antenna to get where 100 Ω input impedance will lie so that we can plot the feed line easily, otherwise, we have to tune randomly on the width side.

Now substituting the values and calculating the y_0

$$100 = 243 \cos^2(\pi y_0/1.04)$$
$$y_0 = 0.3 - 0.4 \, \text{mm}$$

Therefore, the cut size is (0.4×0.4) mm^2. Length of the cut size is equal to 0.3–0.4 mm which is calculated by above given formula, to make the construction of the antenna convenient taking the length as 0.4 mm and the width of the cut size is equal to 0.4 mm which is calculated by 'em:talk microstrip line calculator'.

4 Methodology

In this section, the planned antenna's theoretical analysis, software analysis, experimentation, synthesis, characterization and material selection are discussed in detail.

We have employed Ansys HFSS software for simulation of antennas. Ansys HFSS is a simulation tool for high frequency and high speed electronics. It provides rigorous validation, versatile modelling, adaptive solution that puts analyst quality solvers in the hands of the designer. It can be used for highly accurate parasitic extraction, impedance matching networks and transmission path losses. After the antenna's dimension calculation, the structure is manually designed in HFSS software. The FR4 epoxy substrate is mounted over the ground sheet, and then, the rectangular patch is mounted over the substrate, followed by the feed line designing over the substrate. The input port is assigned as a lumped port. Lumped ports excite or terminate passive circuits and antenna devices in terms of S-parameters, such as impedance matching and insertion loss.

For most PCB practices, FR4 epoxy glass substrates are highly in demand. 'FR' stands for 'flame retardant' and '4' represent 'woven reinforced epoxy resin'. Using FR4 instead of expensive 'Polytetrafluoroethylene' (PTFE) based antenna substrates might result in significant cost reductions, which is why FR4 material is selected for designing substrate. The material is extremely economical and has great mechanical qualities. It is ideal for a wide range of electrical component applications after being assembled [26].

For simulation, the analysis setup is provided to the antenna and the solution frequency is given as 45 GHz as this antenna is designed to work up to 45 GHz, followed by the addition of frequency sweep. In general, frequency sweeps are used to describe the time-dependent performance of a model in the non-destructive contortion range. High frequencies are used to imitate quick movement on small timescales, whereas low frequencies are used to represent sluggish movement on long durations or when at rest.

5 Results and Discussion

In this division, the results of the planned antenna are discussed.

The antenna is simulated inside a radiation box with air surrounding it. The simulated $S(1, 1)$ parameter is shown in Fig. 2. The planned antenna is resonating from 38 to 40 GHz and at 45 GHz the return loss value is − 3.18 dB. The voltage-standing wave ratio (VSWR) is an important metric that determines the total sum of power collected in the conflicting path of an antenna's transmission. VSWR is always a positive integer number. The lower VSWR number provides the better match to the feed line, and more power is provided to an antenna. The VSWR of the planned

antenna is 5.52 at 45 GHz as shown in Fig. 3 by which we can get good far-field radiation quality of the planned antennas.

The connection between antenna and amplifier is very important. Amplifier may require different impedance matching, for example, 10 Ω or 100 Ω, depending upon the characteristic of amplifier. Here, the antenna impedance must be matched using a lossless matching network. Lossless matching network generally consists of capacitors and inductors. It may also consist of coaxial line or transmission line but generally resistance is not used [27–30]. Input impedance is the equivalent impedance that is observed by the source across the circuit. The input impedance of the planned antenna is equal to 100 Ω, and the simulated result of input impedance of planned antenna is shown in Fig. 4 which validates our result.
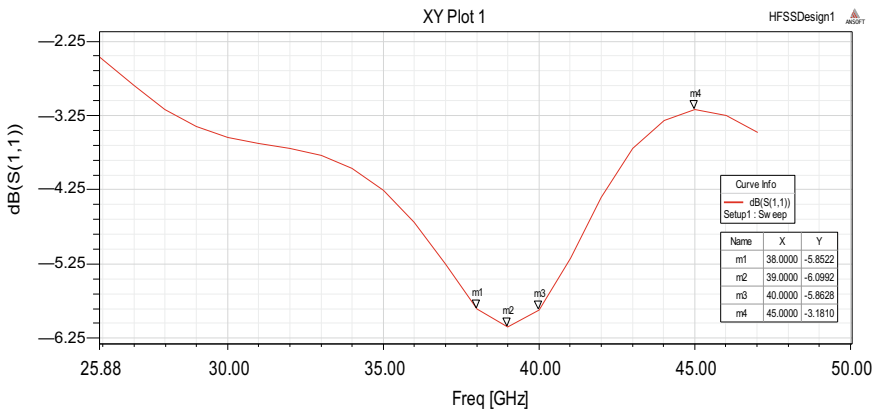
Fig. 2 $S(1, 1)$ parameter of the planned antenna
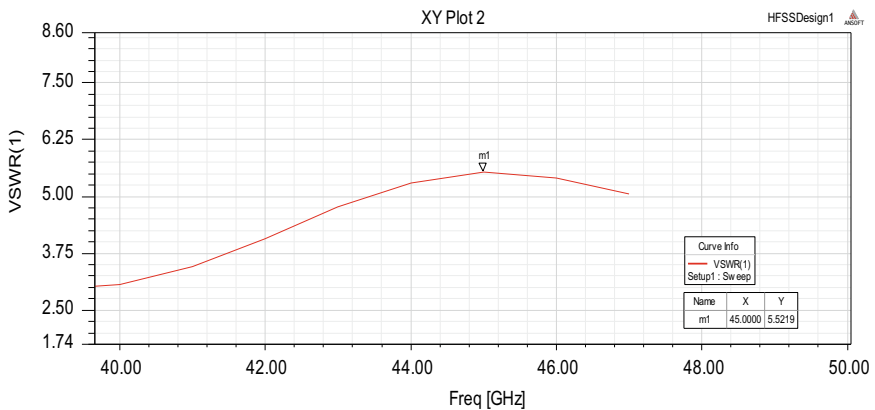
Fig. 3 VSWR of the planned antenna
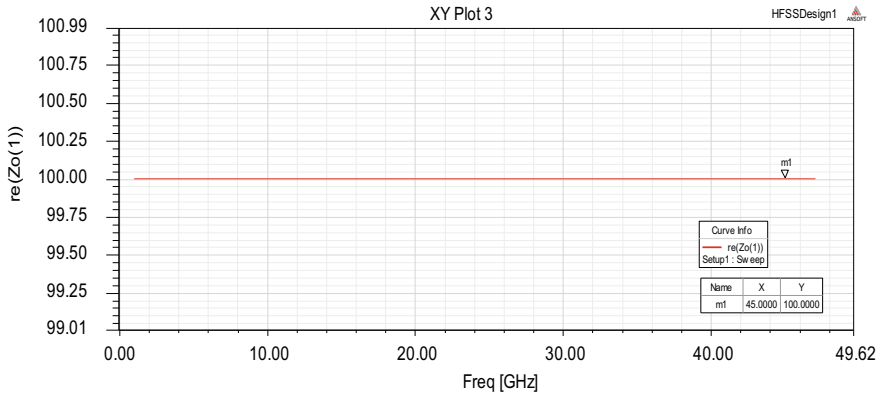
Fig. 4 Input impedance plot of the planned antenna

The simulated 3D radiation model of the planned antenna is revealed in Fig. 5, and the gain plot is revealed in Fig. 6. Gain is one of the most important fundamental parameter of the antenna. Gain and directivity are like two sides of the coin. Directivity is one of the idealistic quantity, whereas gain is something more practical. Directivity is found out at the radiating terminal of the antenna, whereas gain is found at the input terminal of the antenna. In the planned antenna, the maximum gain is in Z-direction which is equal to 3.94 dB and the efficiency of the antenna is 86.61%, i.e. much better in comparison with Refs. [22–25].

Radiation pattern is also the most important parameter of the antenna. These radiations are more in a particular direction and comparatively less in rest of the directions. This property is known as radiation property of an antenna. Figure 7a and b depicts the radiation pattern in H-plane and E-plane of the planned antenna. The radiation pattern is omnidirectional and maximum radiation is in the Z-axis direction.

Bandwidth of the antenna is defined in many ways; it is also related with the gain. Bandwidth of the antenna over a certain band is the gain changes by maximum of 1 dB

Fig. 5 Radiation pattern of the planned antenna
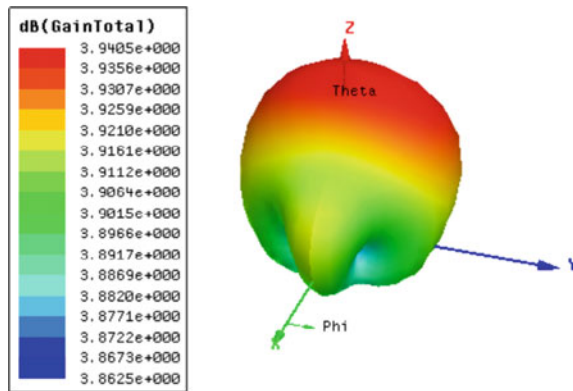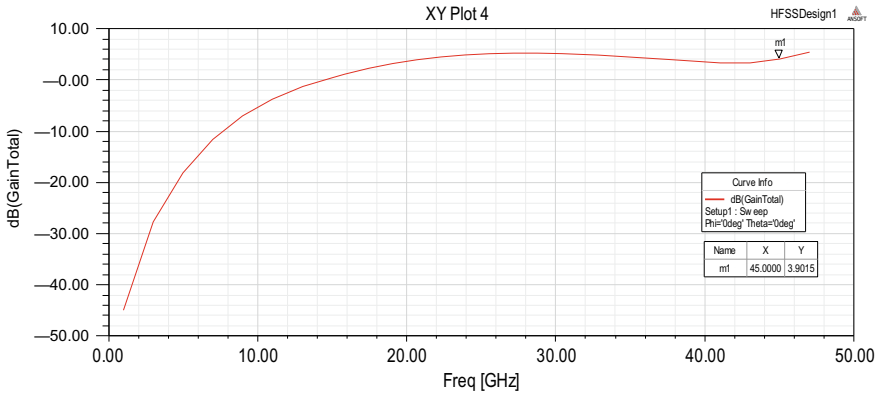
Fig. 6 Gain versus frequency plot of the planned antenna
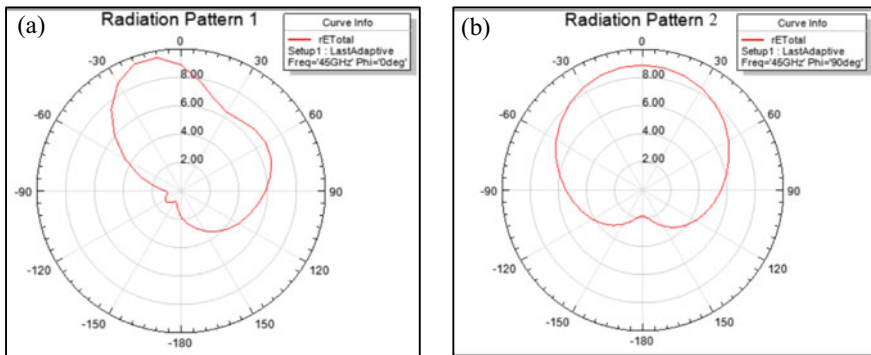
Fig. 7 Simulated radiation pattern in **a** E-plane and **b** H-plane

or 2 dB. It can be defined even for polarization also, a circular polarization bandwidth over which actual ratio is less than 3 dB. Also, the bandwidth of an antenna is the frequency series over which its action complies with a defined standard characteristic [31, 32]. Here, bandwidth of the antenna is defined in stipulations of its reflection coefficient and VSWR. As the planned antenna is designed for input impedance 100 Ω so for this reference the $S(1, 1)$ (dB) versus frequency (GHz) graph should cut at two points at − 5 dB as shown in Fig. 8 and the difference between these two points is the bandwidth of the planned antenna. The simulated bandwidth result is shown in Fig. 8 and the bandwidth of the planned antenna is 4.64 GHz.

The results seem good in comparison with Refs. [22–25]. The comparison chart is revealed in Table 1.
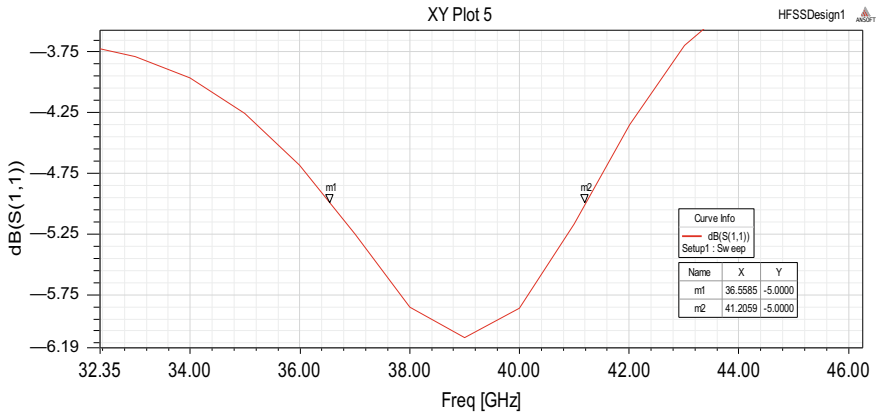
Fig. 8 Simulated bandwidth result of planned antenna

Table 1 Comparative chart

S. No.	Parameters	Antenna				
		mmWave antenna (2014) [22]	Compact microstrip antenna (2019) [23]	Wideband-printed antenna (2020) [24]	Broadband microstrip antenna (2020) [25]	Planned antenna
1	Return loss/$S(1, 1)$	− 42.78 dB at 38.11 GHz	− 20.26 dB at 28 GHz	− 13.35 dB at 3.6 GHz	− 10 dB at 28 GHz	− 3.18 dB at 45 GHz
2	Gain (dB)	9.025	2.48	2.5	5.06	3.90
3	Bandwidth	1.27 GHz	656 MHz	720 MHz	5.57 GHz	4.64 GHz
4	Efficiency	83.30%	Not applicable	Not applicable	80.18%	86.61%
5	Substrate	RT/Duroid 5880	FR4	FR4	RT/Duroid 5880	FR4
6	Working frequency (GHz)	28 and 38	28	3.3–4.2	28	45
7	Simulation software	ADS software	HFSS software	HFSS software	FEKO software	HFSS software

6 Conclusion

In this paper, the high-band microstrip antenna is planned. The antenna is very tiny and can be fitted into small portable devices for the next future 5G communication systems. It is designed to work up to 45 GHz and frequencies can be varying by altering the parameters of the antenna. The antenna shows the omnidirectional radiation pattern for high-band frequency and its resonance frequency can be change or

adjusted easily. The antenna size is reduced, and bandwidth has been enhanced also the efficiency of the antenna is observed 86.61%. The planned antenna has the simple configuration and could be used in various appliances. As the technology moving towards the Internet of Things (IoT), this antenna would help in such technologies.

References

1. Pandey A (2019) Practical microstrip and printed antenna design: Artech House
2. Munson R (1974) Conformal microstrip antennas and microstrip phased arrays. IEEE Trans Antennas Propag 22(1):74–78
3. Garg R (1979) Design equations for coupled microstrip lines. Int J Electron 47(6):587–591
4. Newman E, Tulyathan P (1981) Analysis of microstrip antennas using moment methods. IEEE Trans Antennas Propag 29(1):47–53
5. Pozar D (1982) Input impedance and mutual coupling of rectangular microstrip antennas. IEEE Trans Antennas Propag 30(6):1191–1196
6. Samaras T, Kouloglou A, Sahalos JN (2004) A note on the impedance variation with feed position of a rectangular microstrip-patch antenna. IEEE Antennas Propag Mag 46(2):90–92
7. Geng JP, Jiajing L, Rong-Hong J, Sheng Y, Xianling L et al (2009) The development of curved microstrip antenna with defected ground structure. Prog Electromagn Res 98:53–73
8. Ghassemi N, Mohassel JR, Neshati MH, Tavakoli S, Ghassemi M (2008) A high gain dual stacked aperture coupled microstrip antenna for wideband applications. Prog Electromagn Res B 9:127–135
9. Cao W, Zhang B, Yu T, Li H (2010) A single-feed broadband circular polarized rectangular microstrip antenna with chip-resistor loading. IEEE Antennas Wirel Propag Lett 9:1065–1068
10. Deshmukh AA, Ray KP (2011) Broadband proximity-fed modified rectangular microstrip antennas. IEEE Antennas Propag Mag 53(5):41–56
11. Iwasaki H (1995) A circularly polarized rectangular microstrip antenna using single-fed proximity-coupled method. IEEE Trans Antennas Propag 43(8):895–897
12. Clavin A, Huebner D, Kilburg F (1974) An improved element for use in array antennas. IEEE Trans Antennas Propag 22(4):521–526
13. Mezzavilla M, Zhang M, Polese M, Ford R, Dutta S et al (2018) End-to-end simulation of 5G mmWave networks. IEEE Commun Surv Tutor 20(3):2237–2263
14. Niu Y, Li Y, Jin D, Su L, Vasilakos AV (2015) A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges. Wirel Netw 21(8):2657–2676
15. Sakaguchi K, Haustein T, Barbarossa S, Strinati EC, Clemente A, Destino G et al (2017) Where, when, and how mmWave is used in 5G and beyond. IEICE Trans Electron 100(10):790–808
16. Giordani M, Mezzavilla M, Zorzi M (2016) Initial access in 5G mmWave cellular networks. IEEE Commun Mag 54(11):40–47
17. Syed A, Aldhaheri RW (2016) A very compact and low profile UWB planar antenna with WLAN band rejection. Sci World J
18. Ali MT, Pasya I, Zaharuddin MM, Ya'acob N (2011) E-shape microstrip patch antenna for wideband applications. In: IEEE international RF & microwave conference, pp 439–443
19. Guo J, Yanlin Z, Chao L (2011) Compact broadband crescent moon-shape patch-pair antenna. IEEE Antennas Wirel Propag Lett 10:435–437
20. Patel NH (2020) Design of C-shaped patch antenna for multiband applications. Int J Eng Res Technol 9(6)
21. Kannadhasan S, Shagar A (2017) Design and analysis of U-shaped micro strip patch antenna. In: 2017 third international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB), pp 367–370
22. Chauhan B, Vijay S, Gupta SC (2014) Millimeter-wave mobile communications microstrip antenna for 5G-A future antenna. Int J Comput Appl 99(19):15–18

23. Güneşer TM, Şeker C (2019) Compact microstrip antenna design for 5G communication in millimeter wave at 28 GHz. Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi 12(2):679–686
24. Kapoor A, Mishra R, Kumar P (2020) Compact wideband-printed antenna for sub-6 GHz fifth-generation applications. Int J Smart Sens Intell Syst 13(1):1–10
25. Przesmycki R, Bugaj M, Nowosielski L (2021) Broadband microstrip antenna for 5G wireless systems operating at 28 GHz. Electronics 10(1):1
26. Aguilar JR, Beadle M, Thompson PT, Shelley MW (1998) The microwave and RF characteristics of FR4 substrates. IEE Colloq Low Cost Antenna Technol 2(2)
27. Sahu AK, Misra NK, Mounika K, Sharma PC (2022) Design and performance analysis of MIMO patch antenna using CST microwave studio. Smart systems: innovations in computing. Springer, Singapore, pp 431–441
28. Haykin S, Moher M (2007) Introduction to analog and digital communications. Wiley
29. Yarman SB, Aksen A (1992) An integrated design tool to construct lossless matching networks with mixed lumped and distributed elements. IEEE Trans Circuits Syst I: Fundam Theory Appl 39(9):713–723
30. Kishore K, Akbar SA (2020) Evolution of lock-in amplifier as portable sensor interface platform: a review. IEEE Sens J 20(18):10345–10354
31. Lampard D (1956) Definitions of 'bandwidth' and 'time duration' of signals which are connected by an identity. IRE Trans Circuit Theory 3(4):286–288
32. Amoroso F (1980) The bandwidth of digital data signal. IEEE Commun Mag 18(6):13–24

A Cryptocurrency Price Prediction Study Using Deep Learning and Machine Learning

D. Siddharth and Jitendra Kaushik

Abstract A cryptocurrency is a network-based computerized exchange that makes imitation and double-spending pretty much impossible. Many cryptocurrencies are built on distributed networks based on blockchain technology, which is a distributed ledger enforced by a network of computers. Thanks to blockchain technology, transactions are secure, transparent, traceable, and immutable. As a result of these traits, cryptocurrency has increased in popularity, especially in the financial industry. This research looks at a few of the most popular and successful deep learning algorithms for predicting bitcoin prices. LSTM and Random Forest outperform our generalized regression neural architecture benchmarking system in terms of prediction. Bitcoin and Ethereum are the only cryptocurrencies supported. The approach can be used to calculate the value of a number of different cryptocurrencies.

Keywords Bitcoin · Ethereum · Long short-term memory · Random forest · Time series · Cryptocurrency · Price prediction

1 Introduction

Nowadays, there are over 5000 cryptocurrencies available over the world. However, there are several issues to deal with regarding scientific research. Similarly, they are not highly ranked for market capitalization as market drivers. Another top-ranked pre-mined currency has a third characteristic that cannot be quickly asserted for society's open-sourced non-mineable coins. A controlled blockchain also supports non-mineable coin transactions. We will now see the top-rated two cryptocurrencies up to the market date: bitcoin (BTC) and Ethereum (ETH). These currencies are

D. Siddharth · J. Kaushik (✉)
Department of Data Science, CHRIST (Deemed to be University), Bengaluru, India
e-mail: Jitendrakaushik1986@gmail.com

D. Siddharth
e-mail: d.siddharth@science.christuniversity.in

J. Kaushik
MIT-ADT University, Pune, Maharashtra, India

used for trading and got a lot of profit. That currency they can buy or sell in the coin switch platform. They are network-based exchange mediums, and they use cryptography algorithms to secure the transaction. They have their wallet to exchange all their currency and have their savings. And cryptocurrency is relevant to the stock market in that they buy and sell the shareholders. In this also similar, they will buy and sell currency. But its price fluctuates a lot depending on the parameter we discussed before. Many academics have worked with machine learning and deep learning techniques as well as well-known cryptocurrencies such as bitcoin and Ethereum, as per the literature. Many digital currencies, including TRON, bitcoin, Stellar, and others, have a wide range of applications in banking firms.

The trading companies demand cryptocurrency price prediction because they have to fix the target. So it will get floated like ups and downs. So by a projection of the price, it would be helpful to them to know the strategy and move according to that. So this literature would be beneficial to them.

2   Related Work

Patel et al. [1] in their work, they have highlighted how time series models such as ARIMA, SARIMA, ARCH, and GARCH are widely used to investigate various benefits of implementing. They did, however, rely heavily on the time series forecasting method, which has apparent limitations to assumptions. As per the researchers, neural networks have shown promising results in time series data prediction. They've presented a technique for predicting bitcoin prices incorporating GRU and LSTM models.

Ji et al. [2] only bitcoin estimation techniques based upon bitcoin blockchain data were expected in this study. DNN, LSTM models, CNN, and deep residual networks are just a few of the deep learning techniques they've researched (ResNet). By a bit of margin, LSTM has outpaced models. All deep learning approaches tend to predict the price movement quite well in respect of logistic regression.

Lahiri and Bekiros [3] in their work, they use deep learning to predict virtual currency prices. The research of unpredictable and irregular movements is vital to the predictability of nonlinear systems. They have added to the tectonophysics literature by exploring the complicated properties of the three most widely traded digital currencies. Cryptocurrencies showed substantial self-similarity in training and testing subsamples when this segmentation was performed during deep learning processing.

Livieris et al. [4] this study explains how to estimate cryptocurrency prices and movement using advanced deep learning models and ensemble learning strategies. Apart from that, it resembles deep learning and neural network models.

Phaladisailoed and Numnonda [5] in this investigation, the GRU findings show the best accuracy, although they require more time to calculate than Huber regression. Open, Close, High, and Low may not be sufficient to predict bitcoin values since different aspects, including social sites reactions, rules, and legislation, require them to make accurate predictions.

Aggarwal et al. [6] this paper used RMSE values to compare the various deep learning models on a comprehensive analysis of multiple parameters affecting bitcoin price prediction. The findings reveal that the different deep learning models can adequately predict bitcoin prices. When a good tweet in bitcoin is written, the data of bitcoin price prediction using Twitter sentimental analysis demonstrate a positive link and are presented.

3 Proposed Work

3.1 Data

All data used in this research are everyday records in US dollars for bitcoin, Ethereum, and Ripple, the three most valuable cryptocurrencies by market valuation. From January 1, 2017 until October 31, 2020, all bitcoin data were obtained from https:// www.cryptodatadownload.com. The bitcoin data was divided into three sets for assessment: training, validation, and testing. The training set contained everyday period of January 1, 2017, to February 1, 2018 (1381 data points), and the validation set included data from March 1, 2020, to May 31, 2020 (94 data points). Data from June 1, 2020, to October 31, 2020 (152 data points), were included in the testing set, assuring a significant number of observed and in pieces of data during testing.

- Price: The day's average cryptocurrency price.
- Close: The price of cryptocurrencies at the end of the day.
- Open: The day's starting price of cryptocurrencies.
- Low: Today's bitcoin price is the lowest it's ever been.
- High: Cryptocurrency's highest price of the day.
- Volume: The number of cryptocurrencies that were exchanged in a single day.

3.2 Model Description

Random Forest
Random Forest is a standard bagging method that involves training decision trees in parallel and aggregating the results. Decision trees are very dependent on the data used to train them, and the results can be drastically different when the training data is changed. As a result, the projections are constantly correct. Furthermore, training decision trees is computationally expensive. Because they can't go backward after a split, it has a great danger of overfitting, and they tend to find local optimization rather than global optimization.

LSTM
The RNN version of long short-term memory (LSTM) is smart enough to learn long-term connections. The structure of LSTMs is very comparable to those of RNNs.

However, the repeating unit is quite different. They feature four neural network layers that interact with each other rather than just one.

Scikit-learn and Keras were utilized in this study for data analysis, deep learning, and machine learning models. The TensorFlow software is also employed to create data flow diagrams in this study.

Scikit-Learn

Scikit-learn is a free and open-source data mining analysis package. Python is used to analyze and build machine learning algorithms such as classification, regression, and clustering. In addition to normalization, standardization, and cleaning aberrant or missing data, Scikit-learn may be used to analyze data in a variety of ways.

TensorFlow

TensorFlow is a deep learning framework designed by Google as an open-source project. Dealing with a large number of graphical processing units (GPUs) may be used to train and forecast neural network (NN) architectures, allowing for the adoption of intense deep learning and NN methodologies. Speech recognition, computer vision, robotics, and other sectors might benefit from this design. When graphs are made up of node groups, TensorFlow may construct data flow graphs for processing.

Keras

Keras is a high-level NN open-source library. It is a Python API for neural network programming. Tensorflow, CNTK, and Theano libraries are also supported. Keras can generate models for machine learning, NN, and deep learning. Keras is easy to develop and comprehend, thanks to the division of programs into pieces. The most prevalent creation model components are neural layers, cost functions, optimizers, and activation functions. Python may be used to quickly construct new specified functions or classes.

4 Model Training and Validation

The validation split option in the Keras library's fit function is used to validate the model. The validation split is set at 0.2, a fraction between 0 and 1. The model will separate this portion of the training data from the rest and not be trained on the same data. The model will evaluate the loss and any model metrics on the data set aside at the end of each epoch. Epochs are fine-tuned using various models and values from the SK learn library's grid search csv, with the best value picked. Hyperparameters like as the number of hidden layers, batch size, drop-out rates, and learning rate are all changed on different architectures to produce the best results. The random searching and grid research functions in the SK learn library are used to find the best values.

The error between the output and the supplied goal value is calculated using the loss function. It calculates our distance from the goal value. Many loss functions are

dependent on the problem; thus, we must choose them carefully. The binary pass loss function is utilized in binary classification issues. It is based on the entropy approach, which demonstrates chaos or uncertainty. It is calculated using a probability distribution for a random variable X. A higher entropy value for a probability distribution indicates a more extended time until it is issued. On the other hand, a smaller value indicates a more certain distribution.

The first layer has 40 neurons in the first model, glorot uniform as the kernel initializer and tanh as the activation function. It is reduced to 20 neurons in the second layer, with tanh as an activation function. Drop out is added at a rate of 0.3 between these two layers to add regularization to the model and prevent overfitting. The model is built using the binary cross-entropy loss function and Adam optimizer, with an ideal learning rate of 0.005.

The second model contains 40 neurons in the first layer, with the uniform as the kernel initializer and ReLU as the activation function. It is reduced to 20 neurons in the second layer, and ReLU is used as an activation function. Drop out is added with 0.2 between these two layers to regularize the model and prevent overfitting. It is reduced to 10 neurons in the third layer, and ReLU is used as an activation function. As the research objective is the classification of consumers into churners and non-churners, the output layer contains a component of sigmoid activation. The model was created using the binary cross-entropy gradient descent and the Adams algorithm, and it has an optimum learning rate of 0.01.

5   Results and Discussions

All comparison ensemble methods were tested for machine learning and data mining issues, specifically for forecasting the cryptocurrency price for the next hour (regression) and whether the cost will rise or decrease for the next hour (classification). According to our findings, including deep learning models into an ensemble learning framework improved predictive performance more often than using a single deep learning model. Bagging was the most accurate classification, followed by average and stacking (kNN); stacking (LR) had the best regression accuracy. The confusion matrices showed that the primary learner, stacking (LR), was biassed because the majority of the cases were mistakenly categorized as "Low," whereas bagging and stacking (CNN) exhibited a balanced forecast dispersion among "Lower" and "High," correspondingly. Predictions of "down" or "up." Since bagging can be understood as a perturbation approach to boost resilience, particularly against outliers and very volatile pricing, it's worth noting. The average classifier model based on a modified training data set promotes normalization to such disturbances and better reflects the direction movements of a randomness process represented, according to the numerical analyses.

The model should be tested on new data after it has been trained. The model's performance is evaluated using performance assessment criteria such as accuracy and ROC curve. The plots are drawn between training and testing data, and the measure

utilized is the accuracy of the model's predictions. The loss of training and testing data is depicted in Fig. 1. The cost of training begins at 40,000 and steadily reduces to 8000 as the number of epochs increases from 0 to 100.

The ROC curve of the classifier model is represented in Fig. 2. The ROC curve is a tool for evaluating binary classification issues. The mean square error is plotted on a probability curve. The epochs values are shown on the ROC curve.

For the first model, after training, the model is evaluated on unseen data. Both actual and prediction look similar. It shows models have predicted well. Figures 3 and 4 show the actual price and predicted price inline plot graphs (Fig. 3).
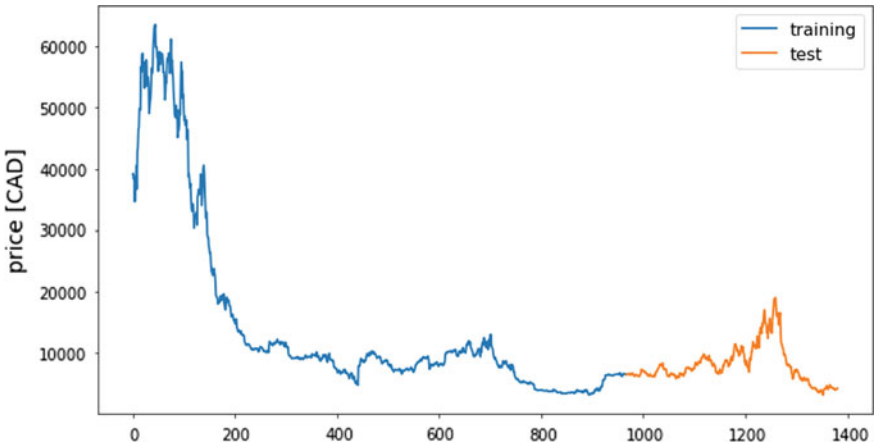
Fig. 1 Training and testing data
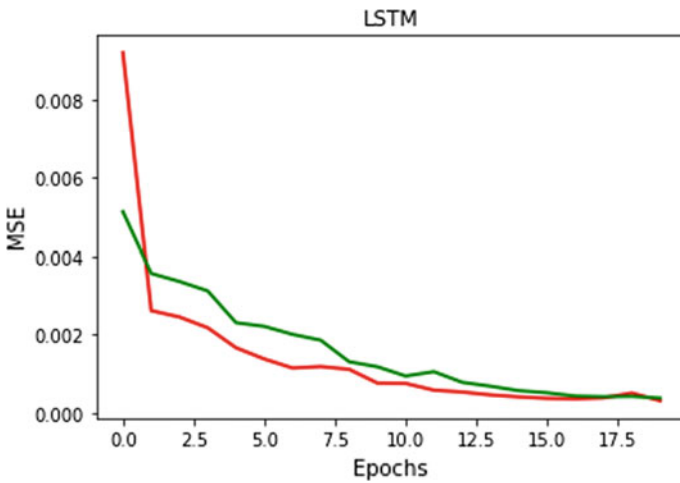
Fig. 2 ROC curve
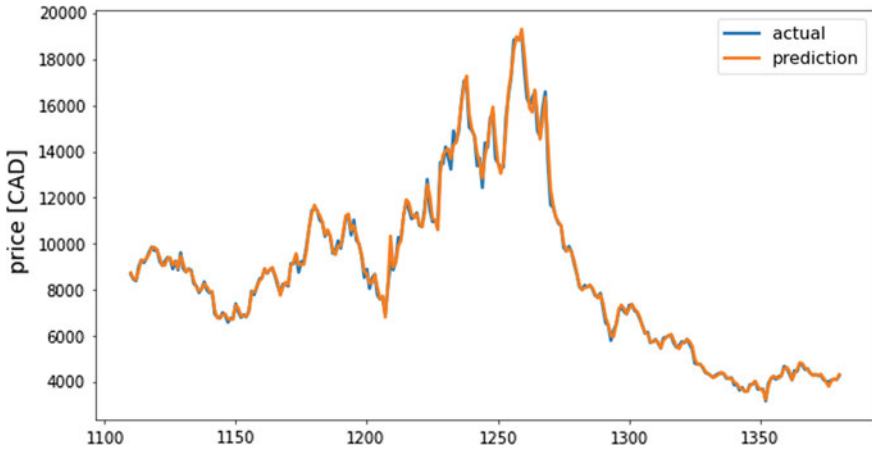
Fig. 3 Actual price and predicted price of LSTM

Fig. 4 Actual price and predicted price of random forest
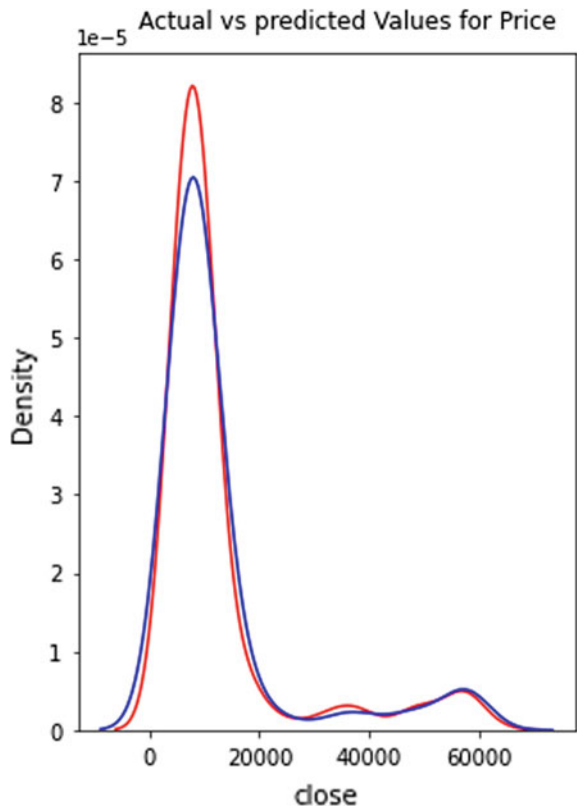
Fig. 5 Actual price and
predicted price of model-2

	Actual	Predicted
654	11396.08	11418.75343
974	6465.12	6485.88480
445	8033.31	8424.61699
609	8063.73	8036.42211
459	9936.40	9845.11827
...	...	...
358	9666.24	9563.78123
724	8115.82	8509.32872
263	10256.20	10165.33374
511	6965.71	7038.67325
359	9518.04	9652.20406

415 rows × 2 columns

Mean squared error (MSE) and *R*-square are two of the most often used metrics to measure the correctness of the data set (R^2). Table 1 shows the MSE and R^2 of all of the algorithms we used. The results indicate that regression models based on deep learning are effective. Such as Random Forest and LSTM outperform Theil-Sen and Huber regression. MSE is 0.01292, and R^2 is 0.9722 or 97.2%, with the best results coming from LSTM. Compared to LSTM and Random Forest, Huber regression takes substantially less time to calculate. Random Forest models are overfitting after cross-validation; also, accuracy is the same.

Table 1 Comparison of models

Models	Mean absolute error	Mean squared error	Root mean squared error	Accuracy (%)
LSTM	0.012922516470067806	0.00048548876771886	0.9722279097949948	97
Random forest	242.38503812047594	322,996.65109717037	568.3279432661835	98.43

6 Conclusion

This paper predicts bitcoin price and movement and developed a deep convolutional neural network based on the multi-structure. The proposed prediction model takes cryptocurrency data as an input and processes it separately, allowing each coin to be exploited and treated independently at first. Each cryptocurrency's data was compiled of inputs to numerous convolutional and LSTM layers used to train the internal structure and discover short- and long-term relationships for each coin. The algorithm then integrates and analyzes the LSTM layers' input vectors' recorded information to produce the final forecast. It's worth noting that, to satisfy the stationarity property, all bitcoin time series were transformed using the returns transformation. Deep learning-based regression models such as Random Forest and LSTM, according to the findings, are more effective than traditional regression models. They outperform Theil-Sen and Huber, regression models. MSE is 0.01292, and R^2 is 0.9722 or 97.2%, with the best results coming from LSTM. Compared to LSTM and Random Forest, Huber regression takes substantially less time to calculate.

References

1. Patel MM, Tanwar S, Gupta R, Kumar N (2020) A deep learning-based cryptocurrency price prediction scheme for financial institutions. J Inf Secur Appl
2. Ji S, Kim J, Im H (2019) A comparative study of bitcoin price prediction using deep learning. Mathematics 7(10):898
3. Lahmiri S, Bekiros S (2019) Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos Solitons Fractals 118:35–40. https://doi.org/10.1016/j.chaos.2018.11.014
4. Livieris IE, Pintelas E, Stavroyiannis S, Pintelas P (2020) Ensemble deep learning models for forecasting cryptocurrency time-series. Algorithms
5. Phaladisailoed T, Numnonda T (2018) Machine learning models comparison for bitcoin price prediction. In: 2018 10th international conference on information technology and electrical engineering
6. Aggarwal A, Gupta I, Garg N, Goel A (2019) Deep learning approach to determine the impact of socio-economic

Regression for Predicting COVID-19 Infection Possibility Based on Underlying Cardiovascular Disease: A Medical Score-Based Approach

Adwitiya Mukhopadhyay and Swathi Srinivas

Abstract The appearance of a novel coronavirus (COVID-19) has presented an immense challenge for the healthcare community around the world. Many patients with COVID-19 have primary cardiovascular (CV) sickness or create intense heart injury throughout the infection. These patients are at exceptionally great danger from COVID-19 because of their fragility and powerlessness for a myocardial involvement. Good comprehension of the exchange between COVID-19 and CV illness is required for these patients' ideal administration. As a growing range of applications for patient management and system incorporation in real time is available, artificial intelligence (AI) can play a decisive role in the emergency department (ED), in fields such as intelligent monitoring, the estimation of clinical results, and resource planning. The proposed system aims to develop an adaptation of a smart medical evaluation method to decide if people with an underlying cardiovascular health disorder would contract COVID-19 based on the limited range of pre-selected variables deemed scientifically necessary and easily calculated when designing clinical judgment regulations.

Keywords Analysis · c-Index · COVID-19 · Logistic regression · Machine learning

1 Introduction

The World Health Organization announced a coronavirus (COVID-19) pandemic on March 11, 2020. Even during an asymptomatic phase, the rapid spread of the virus through regions led to a pandemic because of it is high capability of infection and its capability to propagate during relatively low virulence [1].

A. Mukhopadhyay (✉) · S. Srinivas
Department of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Karnataka, Mysuru, India
e-mail: adwitiya@my.amrita.edu

Cardiovascular patients are at risk of mortality due to their fragility and significance for myocardial involvement. Most COVID-19 patients experience or inflict severe cardiac failure during infection from primary coronary disorders [2]. Administered care is required for these patients. The study findings that conflict with the mortality of COVID-19 and the pre-present comorbidities are responsible for a widespread gap [3]. However, as with other respiratory tract infections, pre-existing coronary conditions and cardiovascular risk factors raise COVID-19 susceptibility [4, 5].

A much wider review of the Chinese Center for Disease Control and Prevention has recorded those patients with elevated blood pressure, diabetes, and CVD; the overall fatality rate (CFR) was 2.3% but extremely high in all patients (6%, 7.3%, and 10.5%, respectively) [1, 6]. Nonetheless, it is evident that regions with the highest death rates, including India, the U.S., Europe, and China, face the most significant pressure on such pre-present continuing circumstances. The reason is, the strain, SARS-CoV-2, the COVID-19-causing agent interacts with the cellular network of cells that express themselves within coronary cardiovascular heart, neuronal and pulmonary alveolar type II cells, an agent of angiotensin changing enzyme 2 (ACE2) [7, 8].

2   Underlying Cardiovascular Disease Impact

Machine learning is an essential platform for prognosis in an industry where patients' survival rate is expected. The limitations of the computational methodology and accessibility issues usually restrict models to a limited range of pre-selected variables. This approach relies on alternative statistical variables. In an emergency room, the machine utilizes the same critical variables and then predicts patient criticality. As a growing range of applications for patient management and system incorporation in real time are available, artificial intelligence can play a decisive role in the emergency room, in fields such as intelligent monitoring, the estimation of clinical results, and resource planning. Similarly, in this instance, the following are the few parameters that play a major role in revealing the underlying disease and help in predicting the likelihood of contracting COVID-19.

2.1 Age

The immune system is a diverse array of cells and molecules that work together to defend, recover us from diseases. In certain areas of the immune system, it is difficult to create an effective immune response to infection as we grow older. The incidence and implications for COVID-19 depend mainly on the patient's age. For adults above 65 years, the risk of mortality is 80% and for adults below 65 years 23 times the risk of death. Comorbidities increase the risk of lethal disease. The molecular differences between young people, middle-aged, and older people were introduced by Amber

L. Mueller, Maesve St. McNamara, and David A. Sinclair discuss that some people suffer mild disorder from COVID-19, but others are life threatening [9].

Age is the most crucial risk factor for coronary, with around threefold chances each decade of life. Adolescents can begin to form coronary fatty strips. 82% of those who suffer from coronary heart disease are reported to be 65 and older. At the same time, every decade after the age of 55, the chance of stroke doubles. The leading cause for this is blood vessels also age as the age increases. They become inflexible when they age, and it becomes more difficult for blood to travel rapidly. While the new COVID-19 can lead to death for adolescents too, it causes the most serious health problems for adults over the age of 60—particularly lethal for those aged 80 and older [10]. This is primarily attributed to the number of chronic health problems in older adults. Often, as people grow older, their immune system loses its strength gradually, suggesting that they are more vulnerable to some type of infection, especially COVID-19.

2.2 Gender

This is partly due to the number of chronic conditions of older people. Conditions like asthma, heart disease, and other chronic illnesses can contribute to more serious symptoms and complications. Sometimes, as individuals get older, their immune system is becoming increasingly less heavily affected by some form of virus, in particular a new one, like COVID-19. The incidence of COVID-19 is more prevalent in men and women with low outcomes and mortality, irrespective of age [11, 12].

2.3 Diabetes

Diabetes is a big risk factor for women to develop cardiovascular diseases. Diabetic women may have very slight signs of heart failure. Blood vessels and nerves will suffer from increased blood glucose diabetes over time. Adults with diabetes are the most frequent cause of death, and mortality rate is double compared to non-diabetic, as cardiovascular disease are highly experienced by them [13, 14].

Type 1 diabetes is a disorder of the autoimmune system that causes the pancreas' beta cells to be destroyed. The hormone insulin is created little to no. ketones are chemicals formed by the body as fat is decomposed by energy. This happens if the body has not adequate insulin. They increase acidity, which can be very harmful, as ketones accumulate in the blood.

In Type 2, because of an insulin resistance issue, the body does not produce adequate insulin or use any available insulin effectively. Diabetes from conception develops during infancy and is usually removed after pregnancy. That being said, it is more likely that individuals with gestational diabetes will experience type 2 diabetes

in later life. Gestational diabetes is also highly dangerous for a person who develops COVID-19 [14, 15].

Based on the current understanding, type 1 or gestational diabetes will raise the risk of serious COVID-19 disease [16]. As the researchers have explained, diabetes in people infected with the COVID-19 has long been known to increase the risk of catastrophic COVID-19 significantly. According to researchers led by Dr. Yang Jin at Union Hospital and Tongji Medical College in Wuhan, China, elevated blood sugar (glucose) was also associated with more serious diseases and complications, testing at the time of hospital entry. The researchers reported on 10 July in the journal Diabetology elevated levels of fasting blood glucose were highly predictive of death [16]. In a press release of the European Organization for the Study of Diabetes, the odds for COVID-19 harmful complications were both four times higher for individuals in the elevated blood sugar community and two and six times higher than for those with pre-diabetes blood sugar level [17]. The team said that the impact of high blood sugar on mortality risk is irrespective of whether a patient has more or less serious pneumonia, "regardless of whether the patient has COVID-19 linked pneumonia" [16].

2.4 Blood Pressure

Blood pressure represents the push of blood against artery walls. Each time the heartbeats, it pumps blood into the arteries, resulting in the highest blood pressure as the heart contracts. With elevated blood pressure, arteries may have enhanced blood flow resistance, making it difficult for the heart to pump blood and can cause damage to the heart's arteries over time. The new research indicates that unregulated or untreated elevated blood pressure individuals might be at the possibility of serious illness with COVID-19 [18, 19]. Early data analysis both from China and the U.S. indicates that the most common pre-existing condition among hospitalized patients is high blood pressure which affects 30–50% of the patients [18].

The higher risk of COVID-19 in high blood pressure patients' is due to a weaker immune system. The immune system is less likely to suppress the virus because of long-term health problems and aging. Almost two-thirds of those aged 60 suffer from elevated blood pressure. The higher risk is, however, not the elevated blood pressure itself, but some medicines that are used to treat [20].

COVID-19 may also directly harm the heart, which is highly dangerous if the heart is already affably damaged by hypertension. The virus will cause the heart muscle inflammation known as myocarditis, making pumping harder for the heart. Past findings have found people with coronary disease who are at a greater risk of a heart attack with respiratory disease like influenza or earlier COVID-19 [21, 22].

2.5 Cholesterol

High cholesterol ("bad" cholesterol) is expected to contain high levels of low-density lipoprotein (LDL). The chance of a heart attack is also raised by elevated levels of triglyceride, a blood fat form that is linked to the diet. However, the risk of heart disease is minimized by elevated levels of HDL cholesterol ("good" cholesterol). COVID-19 is a lipid surface RNA virus. Thus, the assembly, proliferation, and infectiveness of these viral particles are dominated by the cholesterol biosynthesis pathways. The drug modification for cholesterol, particularly statins, has been antivirally hypothesized. Such medicines lower synthesis, systematic cholesterol uptake, or have clear antiviral effects that change the cholesterol goal cell membrane. Further pleiotropic symptoms are associated with non-lipid-like statins. This involves enhanced endothelial activity, the stability of an atherosclerotic plate, anti-inflammatory effect, and immunomodulation. This additional statin property could have potential benefits for COVID-19 patients [23, 24].

2.6 Obesity

Obesity biology contains compromised immunity, systemic inflammation, and blood which can exacerbate COVID-19, and because obesity is so stigmatized, people may escape medical treatment for obesity [25].

Obesity is more likely in individuals with other independent COVID-19 risk factors than normal weight people. They are also vulnerable to a metabolic syndrome of blood sugar, weight, or both unhealthy and elevated blood pressure. Fat in the abdomen presses on the diaphragm and allows the large muscle that lies under the chest to affect the lungs and reduce air fill. This lowered amount of lung results in the airways closing in the bottom lungs, with more blood coming to oxygen than in the top lungs. If already with such a malfunction, it gets worse more quickly of COVID-19. These mechanical difficulties are exacerbated by other concerns. Second, the blood of obese individuals is more likely to clot. Beverley Hunt, a scientist who is a specialist in blood clotting at Guy's and St. Thomas' hospitals in London said, the endothelial cells that line the blood vessels usually not to clot. However, COVID-19 changes signaling since the virus injures endothelial cells that respond by stimulating the mechanism of clotting. The chance of coagulation increases.

Immunity is also deteriorating in obese people in part because fat cells invade the organs in which immune cells are developed and processed, says Fairfield University nutritionist Catherine Andersen [25]. The immune system lacks its tissue which makes it less efficacious to either protect the body from infections or respond to a vaccination. They lose tissue against adipose. The issue is not only fewer immune cells but also poorer immune cells [26, 27].

3   Methodology

The proposed system is trained from the medical dataset to predict the contribution of underlying cardiovascular disease in contracting COVID-19 along with the proposed medical scoring chart using the essential vitals alone. The architecture is shown in Fig. 1.

3.1   Risk Score Module

In this module, each vital parameter is analyzed based on the proposed scoring chart, and a risk score is assigned. These parameters play a major role in detecting the likelihood of a patient catching COVID-19, and these parameters are individually analyzed [28]. Each parameter's analysis is shown in Tables 1, 2, 3, 4 and 5.

3.2   Cardiovascular Module

The prognosis model is trained using logistic regression as it is a classification algorithm which is used to find the likelihood of a successful occurrence and a failure of the event. The dependent variable is used if it is binary in essence. It facilitates the classification of data in discrete groups by examining the relationship from a collection of labeled data. It learns from the given dataset; then, it introduces a nonlinearity in the shape of sigmoid. Strong precision for several single datasets and
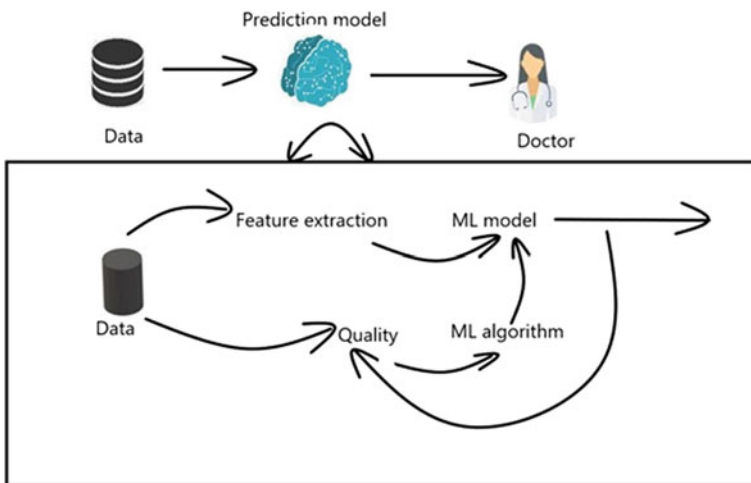
Fig. 1 Architecture diagram of a risk score model

Table 1 Probability and risk score for male and female concerning age

Age	Probability (male) (%)	Probability (female) (%)	Score
< 1	0.013	7.24	2
1–4	6.726	5.792	3
5–14	0.019	0.014	2
15–24	1.785	0.138	1
25–34	0.814	0.547	1
35–44	2.213	1.287	1
45–54	5.960	3.436	1
55–64	14.127	9.323	2
65–74	42.457	18.039	3
75–84	28.211	26.429	4
85+	24.342	40.772	5

Table 2 Probability and risk score for sugar level, BMI, systolic, and temperature

Probability (%)	Range of sugar level	Range of body mass index level	Range of systolic value	Temperature (in Celsius)	Score
20	101–120	18.5–24.9	< 120	34–36	1
30	121–140	25.0–29.9	120–129	< 34 36–37	2
40	141–160	30.0–34.9	130–139	38–39	3
60	161–180	35.0–39.9	140–180	39–40	4
80	> 180	> 40	> 180	40–41	5

Table 3 Probability and risk score for cholesterol

Range of cholesterol level LDL	Range of cholesterol level HDL	Probability	Score
< 100	> 60	Normal	1
100–129	< 50 (F)	Elevated	3
130–159	> 40 (M)	High	4
> 160		Very high	5

Table 4 Risk score for pulse-oxy level

Range of oxygen level (mmHg)	Probability	Score
90–120 (75–100 mmHg)	5%	2
60–90 (< 75 mmHg)	Hypoxemia (organ failure and cardiac arrest)	3
> 120	Hyperoxemia	4
< 60 mmHg	Supplement needed	5

Table 5 Impact level based on risk

Total score	Impact level range (%)	Impact level
> 40	> 50	5
30–40	30–49	4
20–30	20–29	3
10–20	10–19	2
0–10	< 10	1
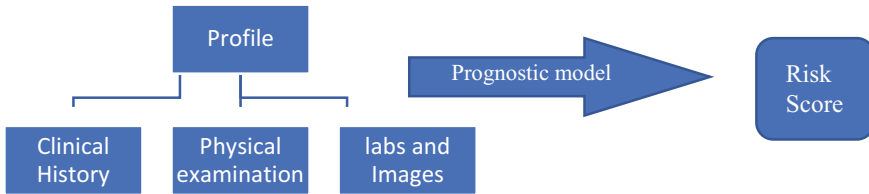
Fig. 2 Architecture diagram of prognosis model

fits well with the linear isolation of the dataset. The model coefficients can be viewed as measures of functionality. The logistic regression is less likely to bypass but can be overridden in large datasets. Regularization strategies are considered to ensure that these scenarios are not overfitted [29].

Prognostic models are a system that inputs and provides a risk score for that patient in a patient's profile. The profile of the patient will contain health reports and all past treatments and major diseases. The prediction model will assign the patient a risk score. Today, risk values can be arbitrary numbers or likelihoods. The predictive model informs about the coefficients or weights of each function. Architecture of prognosis model is represented in Fig. 2.

For example, if the patient has hypertension and diabetes and their age is multiplied by the associated value. The first attribute hypertension is not only a laboratory value, but the lateral log is considered. Likewise, for the remaining two laboratory values, until they are plugged into a model, the normal log of the values is taken. Assume that the relation of risk to characteristics is linear in the natural log of features, which is usually instead of the features themselves. Secondly, in intercept function, the intercept value is always 1, so the intercept is the estimated risk value if all the values are 0. To calculate the patient's score, the summation of product of natural log of the function and coefficient is taken. The other functions of the lab are intercept. For convention, this score is multiplied by 10 and rounded to get an output MELD score of 10 refer Eq. 1.

$$\text{Score} = \sum \text{Coefficient} \times \log(\text{value}) \tag{1}$$

The score tells us the insightful relative to the score of other patients. In this proposed model, cardiovascular risk is predicted by introducing a couple of features. In addition to the natural log of lab value traits, the product of two is taken as one characteristic. These are referred to as interaction terms. Second, a negative coefficient is correlated with such traits. The negative coefficient implies that this characteristic contributes negatively, which decreases the ranking. This makes sense as they are good amount of vital parameter. The feature, multiplied by the associated coefficient and repeated for the other functions, summarizes all contributions and gives a risk equation. The danger equation should not be linear in its characteristics, but linear in a natural log or log base 10 of the characteristics. Finally, risk equations can also contain concepts of interaction refer Eq. 2.

$$\text{Score} = \sum \begin{cases} \text{Coefficient} \times \log(\text{value}) \\ +\text{Coefficient} \times \log(\text{interaction terms}) \end{cases} \tag{2}$$

This sum is converted to a risk score using Eq. 3.

$$\text{Risk} = 1 - 0.9533^{e^{\text{sum}-86.61}} \tag{3}$$

Conditions of interaction can detect a dependence between variables. The influence of the two variables is independent without interaction terms. The basic idea behind evaluating the risk model is to compare the risk scores it assigns to pairs of individuals. The model is tested using c-index. The c-index measures the discriminatory power of a risk score. Intuitively, a higher c-index indicates that the model's prediction is in agreement with the actual outcomes of a pair of patients. To evaluate risk scores, knowing whether the patients had the event or not that is within 10 years. If patient 'A' died within the next 10 years, but patient 'B' did not, these numbers did not have to be between zero and one; they do not have to be probabilities, all that wanted from a model is a higher risk that is assigned to patient 'A' then to patient 'B' (refer Table 6).

Table 6 Permissible and non-permissible pairs

Patient 'A'	Patient 'B'	Outcome (death)	Resulting pair
High risk		A	Concordant
	High risk	A	Not concordant
High risk		B	Not concordant
	High risk	B	Concordant
High risk	High risk	A	Ties
High risk	High risk	B	Ties
High risk		A, B	Non-permissible pairs
	High risk	A, B	
High risk	High risk	A, B	

If both patients die in 10 years or both patients do not die in 10 years, these pairs cannot be used to determine who should have a higher risk score. In the evaluation of prognostic models, only, pairs where the outcomes are different are considered. A pair where the outcomes are different is called a permissible pair. With such pairs, prognostic models are evaluated. Hence, prognostic models are evaluated by giving ties half the weight as concordant pairs. The c-index is computed using Eq. 4.

$$c \text{ - index} = \frac{\text{\#concordant pairs} + 0.5 \times \text{\#riskties}}{\text{\#permissible pairs}} \tag{4}$$

c-index has the interpretation of Eq. 5.

$$P(\text{score}(A) > \text{score}(B)|Y_A > Y_B) \tag{5}$$

where A and B are patients, Y_A, Y_B are patient's outcome [30]. The flowchart of the cardiovascular module is represented in Fig. 3.

4 Results

AI model is trained by applying the linear regression on the heart disease dataset from Cleve-land Clinic Foundation, Hungarian Institute of Cardiology, Budapest, V.A. Medical Center, Long Beach, CA, and University Hospital, Zurich, Switzerland. While the databases have 76 raw attributes, only, 8 of them are used, i.e., age, sex, blood pressure, sugar level, smoking habits, cholesterol, blood sugar level, and history of diabetes. On training the above dataset and evaluating using c-index, an accuracy rate of 73% was achieved for the mean imputed dataset. Considering this outcome with the proposed chart's risk score, the likelihood of contracting COVID-19 is predicted.

5 Conclusion

The proposed system aims in calculating the risk score and probability of contracting COVID-19 for a patient with underlying cardiovascular disease. The study is based on the analysis of vital values concerning the recent outcomes and studies in the area of COVID-19 analysis. Various factors like age, gender, blood pressure, sugar level, pulse-oxy level, body mass index, pulse rate, temperature play a major role in identifying the underlying cardiovascular diseases like hypertension, diabetes, obesity, and cholesterol. Keeping the number of vital parameters minimal, on estimating the risk of underlying cardiovascular disease, the likelihood of contracting COVID-19 is predicted. Firstly, the vitals are analyzed using the proposed risk score chart which is based on the limited range of pre-selected variables deemed scientifically necessary
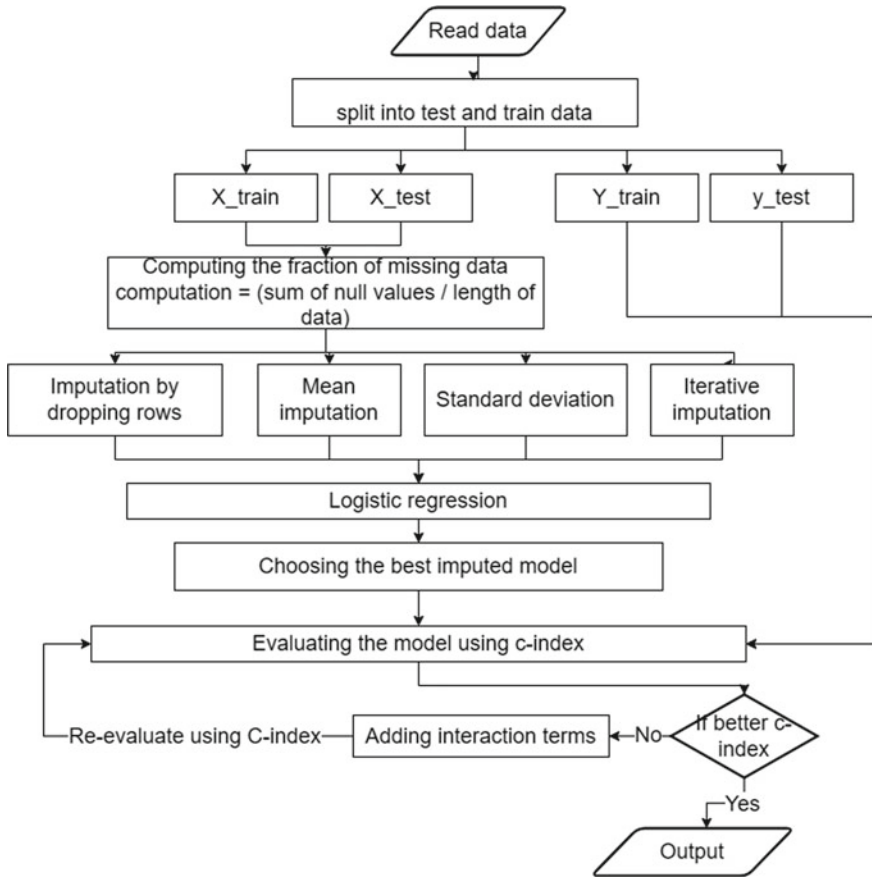
Fig. 3 Flowchart of cardiovascular module

and easily calculated when designing clinical judgment regulations. Secondly, the risk of cardiovascular disease is estimated using the machine learning module, and the overall risk is predicted.

References

1. Bansal M (2020) Cardiovascular disease and COVID-19. Diabetes Metab Syndr 14(3):247–250. https://doi.org/10.1016/j.dsx.2020.03.013
2. Tan W, Aboulhosn J (2020) The cardiovascular burden of coronavirus disease 2019 (COVID-19) with a focus on congenital heart disease. Int J Cardiol 309:70–77. https://doi.org/10.1016/j.ijcard.2020.03.063
3. Ssentongo P, Ssentongo AE, Heilbrunn ES, Ba DM, Chinchilli VM (2020) The association of cardiovascular disease and other pre-existing comorbidities with COVID-19 mortality: a

systematic review and meta-analysis. medRxiv. https://doi.org/10.1101/2020.05.10.20097253

4. Pimentel MAF, Redfern OC, Hatch R, Young JD, Tarassenko L, Watkinson PJ (2020) Trajectories of vital signs in patients with COVID-19. Resuscitation 156:99–106. https://doi.org/10.1016/j.resuscitation.2020.09.002

5. Nishiga M, Wang DW, Han Y, Lewis DB, Wu JC (2020) COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives. Nat Rev Cardiol 17(9):543–558. https://doi.org/10.1038/s41569-020-0413-9

6. Li B et al (2020) Prevalence and impact of cardio-vascular metabolic diseases on COVID-19 in China. Clin Res Cardiol 109(5):531–538. https://doi.org/10.1007/s00392-020-01626-9

7. Madjid M, Safavi Naeini P, Solomon SD, Vardeny O (2020) Potential effects of coronaviruses on the cardiovascular system: a review. JAMA Cardiol 5(7):831–840. https://doi.org/10.1001/jamacardio.2020.1286

8. Driggin E et al (2020) Cardiovascular considerations for patients, health care workers, and health systems during the COVID-19 pandemic. J Am Coll Cardiol 75(18):2352–2371. https://doi.org/10.1016/j.jacc.2020.03.031

9. Mueller AL, Mcnamara MS, Sinclair DA (2020) Aging_COVID19. Aging 12(10):9959–9981

10. Singh HP, Khullar V, Sharma M (2020) Estimating the impact of Covid-19 outbreak on high-risk age group population in India. Augment Hum Res 5(1). https://doi.org/10.1007/s41133-020-00037-9

11. Jin JM et al (2020) Gender differences in patients with COVID-19: focus on severity and mortality. Front Public Health 8:1–6. https://doi.org/10.3389/fpubh.2020.00152

12. Burki T (2020) The indirect impact of COVID-19 on women. Lancet Infect Dis 20(8):904–905. https://doi.org/10.1016/S1473-3099(20)30568-5

13. Gianchandani R et al (2020) Managing hyperglycemia in the covid-19 inflammatory storm. Diabetes 69(10):2048–2053. https://doi.org/10.2337/dbi20-0022

14. Joshi AM, Shukla UP, Mohanty SP (2020) Smart healthcare for diabetes: a COVID-19 perspective. arXiv

15. COVID-19 and diabetes: risks, types, and prevention. https://www.medicalnewstoday.com/articles/covid-19-and-diabetes#types. Accessed 27 Jan 2021

16. Certain medical conditions and risk for severe COVID-19 illness | CDC. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fneed-extra-precautions%2Fgroups-at-higher-risk.html#smoking. Accessed 19 Jan 2021

17. COVID-19 may spike blood sugar, raising death risk. https://www.webmd.com/lung/news/20200713/covid-19-may-spike-blood-sugar-raising-death-risk#1. Accessed 27 Jan 2021

18. Ran J et al (2020) Blood pressure control and ad-verse outcomes of COVID-19 infection in patients with concomitant hypertension in Wuhan, China. Hypertens Res 43(11):1267–1276. https://doi.org/10.1038/s41440-020-00541-w

19. Specialties M (2020) Medication for high blood pressure can improve Covid survival rate, reduce severity of infection, pp 1–9

20. Covid-19 survival rates: medication for high blood pressure can improve Covid survival rate, reduce severity of infection. Health News, ET HealthWorld. https://health.economictimes.indiatimes.com/news/diagnos-tics/medication-for-high-blood-pressure-can-improve-covid-survival-rate-reduce-severity-of-infection/77738021. Accessed 27 Jan 2021

21. COVID-19 and high blood pressure: Am I at risk?—Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/corona-vi-rus/expert-answers/coronavirus-high-blood-pressure/faq-20487663. Accessed 27 Jan 2021

22. High Blood Pressure & Coronavirus (Higher-Risk People): Symptoms, Complications, Treatments. https://www.webmd.com/lung/coronavirus-high-blood-pressure#1. Accessed 27 Jan 2021

23. Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information (2020), pp 19–21

24. Kočar E, Režen T, Rozman D (1866) Cholesterol, lipoproteins, and COVID-19: basic concepts and clinical applications. Biochim Biophys Acta Mol Cell Biol Lipids 2:2021. https://doi.org/10.1016/j.bbalip.2020.158849
25. Wadman M (2020) Why COVID-19 is more deadly in people with obesity—even if they're young. Science. https://doi.org/10.1126/science.abe7010
26. Popkin BM et al (2020) Individuals with obesity and COVID-19: a global perspective on the epidemiology and biological relationships. Obes Rev 21(11). https://doi.org/10.1111/obr.13128
27. Why COVID-19 is more deadly in people with obesity—even if they're young | Science | AAAS. https://www.sciencemag.org/news/2020/09/why-covid-19-more-deadly-people-obesity-even-if-theyre-young. Accessed 20 Feb 2021
28. Mukhopadhyay A, Swathi S, Srinidhi A (2019) Mechanisms for monitoring of blood pressure and analysis of patient criticality level in emergency situations. In: 2019 global conference for advancement in technology, GCAT 2019. https://doi.org/10.1109/GCAT47503.2019.8978463
29. Varun S, Mounika G, Sahoo P, Eswaran K (2019) Efficient system for heart disease prediction by applying logistic regression. Int J Comput Sci Technol (IJCST) 10(1):13–16
30. AI for Medical Prognosis | Coursera. https://www.coursera.org/learn/ai-for-medical-prognosis. Accessed 01 Oct 2021

Evolution of 5G: Security, Emerging Technologies, and Impact

Varun Shukla, Poorvi Gupta, Manoj K. Misra, Ravi Kumar, and Megha Dixit

Abstract The review paper aims to generalize the important facts and issues of Fifth Generation (5G) in the coming future with its emerging technologies. 5G the fifth generation superfast wireless network, which is supervised and engineered for the rapid increase in the network speed. The paper rundowns through the key requisites that are gigabytes per second data rates to end user, increment in number of connected devices and latency rate below per micro second. This paper presents a review on emerging technologies like Massive MIMO technology, Spectrum Sharing (SS), Interference Management (IM), and Device-to-Device communication (D2D). To bring a new technology into reality across the globe, some important factors and challenges like fiber infrastructure, availability of base stations, low data speed, and high rates, etc., are discussed in this paper. The intent of this article is an exclusive study of 5G in all aspects considering security issues and emerging technologies (and their respective impact) like Internet of Things (IoT), Filtered OFDM (f-OFDM), Artificial Intelligence (AI), Virtual Reality (VR), and Augmented Reality (AR), etc.

Keywords Data communication · Device-to-device (D2D) · Fifth generation (5G) · Internet of things (IoT) · Security

Organization of the paper: Introduction is given in Sect. 1. Section 2 talks about customized need (Indian Perspective) and Emerging Technologies. Some challenges in implementation of 5G are discussed in Sect. 3. Advanced technologies used in

V. Shukla · P. Gupta · M. Dixit
Department of Electronics and Communication, Pranveer Singh Institute of Technology, Kanpur, Uttar Pradesh, India

M. K. Misra
Department of Computer Science Engineering, Pranveer Singh Institute of Technology, Kanpur, Uttar Pradesh, India

R. Kumar (✉)
Department of Electronics and Communication, Jaypee University of Engineering and Technology, Guna, Madhya Pradesh, India
e-mail: ravi.kumar6@gmail.com

5G are discussed briefly in Sect. 4. Section 5 is all about the security concerns and conclusion and future scope is given in Sect. 6.

1 Introduction

A 5G technology has become the demand of the present era. It will be the machinery linking billions of devices together. For the effective evolution of Fifth Generation (5G) network, there is a need of an outlay which will help in enhancing the economic benefits of the network. The enormous change in speed of Fourth Generation (4G) to Fifth Generation (5G) and high-intent screens of 5G must guide the regular internet users to engage e-commerce activities and enhance more purchase online. A 5G wireless networks have acquired massive attention in both academic and industry alike which will enable a huge vertical application by associating heterogeneous devices and machines [1, 2]. The challenges faced by the technical officials while developing 5G network after 4G LTE (Long Term Evolution)) are: 1–10 GBPS (Giga Bits Per Second) in real time network, low latency rate i.e. 9–10 ms (milliseconds), high bandwidth & spectrum density, low cost and longer battery life as shown in Fig. 1.

Many explicate methods and customized processes are required for 5G and Internet of Things (IoT) such as Machine-to-Machine (M2M), compatibility with IP addressing and Low-power WAN (LPWAN), etc. In the present smart world, technologies are getting enhanced day by day with a new concepts like Big Data, IoT, M2M, etc. 4G will not be that capable of lifting the requirements mentioned by automation technology. Emergence of Industrial IoT (IIoT) has to be handled very prudently with respect to automated smart control systems [3].

Comparatively 5G consumes more power than 4G and the reason is that 5G works at high speed and low latency. If the device is connected with Wi-Fi then the power consumption can be reduced. Battery consumption is also a subject of particular device and processor. The task of providing a rapid entry to reinforce
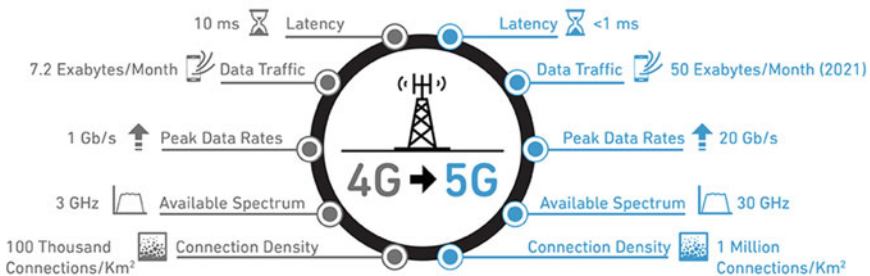
Fig. 1 Comparison of 4G and 5G
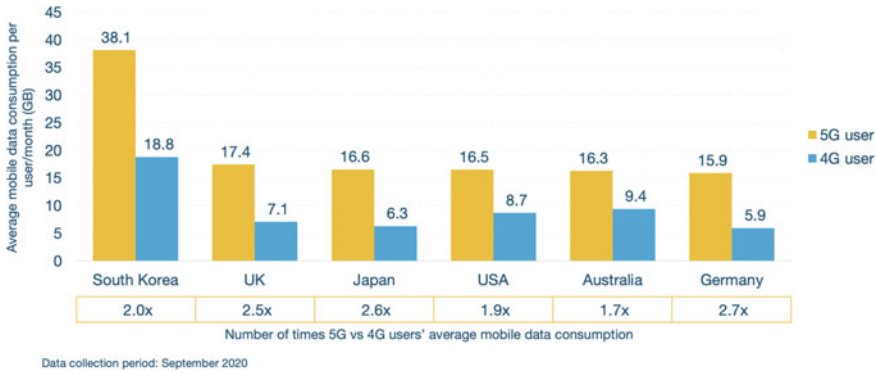
Fig. 2 Comparative data usage of 4G and 5G with respect to various countries

multiple devices for transmitting trivial data is an eminent challenge. The traffic from sensors is latency-sensitive and requires an ultra-resilient communication link. Extreme data rates for the users and subtle latency for the machines have to be supported. It is important to mention that 4G communications rely on Orthogonal Frequency Division Multiplexing (OFDM) [4].

Emerging technologies specified further lead to a synchronized procedure of enduring 5G as the key to make the digital world more convenient considering all aspects of development. These technologies not only enhance the system speed but also help in reduction of the sustaining problems which occur due to the incorporation of new technology into work. 5G will drive the future advancement of the internet. User will be able to download a 15 GB movie file in just 6 s, an autonomous car travelling at up to 100 km per hour will be capable of receiving a stop signal in case of any danger detected with virtually zero latency. Many more users by 2025 are expected to avail internet services in comparison to last 5 years and this tremendous growth will be significant. The data of September 2020 is shown in Fig. 2 and it clearly demonstrates the popularity of 5G over 4G and needless to say that it will bring many opportunities.

It is also interesting to mention that different service providers in India (considering three major market players: Reliance Jio, Airtel, and Vodafone Idea) have different plans for 5G and its related services as shown in Fig. 3 [5].

2 Customized Need (Indian Perspective) and Emerging Technologies

Technology in this modern era keeps on changing in every next generation. As India being a developing country, we need an efficiently successive transformation to lead in the technical world via establishing 5G with a smooth procedure. Considering the present scenario in India, after the utilization of 4G, now people are in full swing to
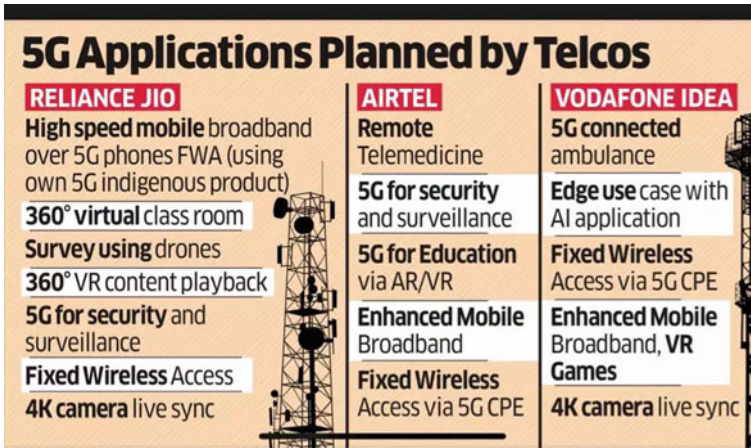
Fig. 3 Comparing the plans of service providers in India

switch to 5G in a progressive manner. 5G not only provides accuracy and improvised speed but includes enormous IoT where billions of devices will be connected at a time on the network with high speed and virtually zero latency. Therefore, presently the main aim is to understand challenges occurring in transformation to 5G [6]. Total 5G connections are expected to 11% by 2025 according to GSMA intelligence, Ericson report as shown in Fig. 4 [7].
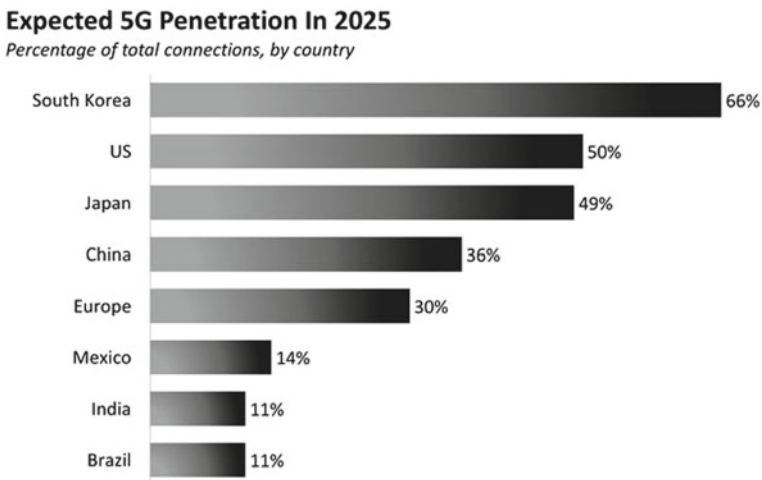
Fig. 4 Expected increment in connections by 2025 with respect to various countries

2.1 Massive MIMO Technology

Massive Multiple Input Multiple Output (MIMO) is an advanced form of MIMO technology as it is a technique of adding a greater number of antennas for multiple output at the base stations which enhances focus energy and hence provides better efficiency. The MIMO technology is a part of LTE-A and is based on the concept of structural multiplexing [8]. Spatial multiplexing is a significant technique of transmitting independent channels separated in space for a more frequent and smooth use of 5G. Massive MIMO has the potential to upgrade the radiated energy efficiency by 100 times and simultaneously increases the capacity of the order of 10 or more [9, 10]. Basically, the radiated energy efficiency from the base antennas is based on the principle of coherent superposition. The expense of transmitted power gets increased as the zero forcing is used to suppress the leftover interference between the terminals [11]. The prime concern of making a consistent and speedy network, reduced latency proves to be a basic demand in the wireless communication whereas massive MIMO permits a prominent decrease in the latency by avoiding fading dips. Massive MIMO has a dominating advantage of low power consumption and less costly components which will help in globalizing 5G especially in developing and under-developed countries [12]. A conceptual illustration for massive MIMO is shown in Fig. 5.
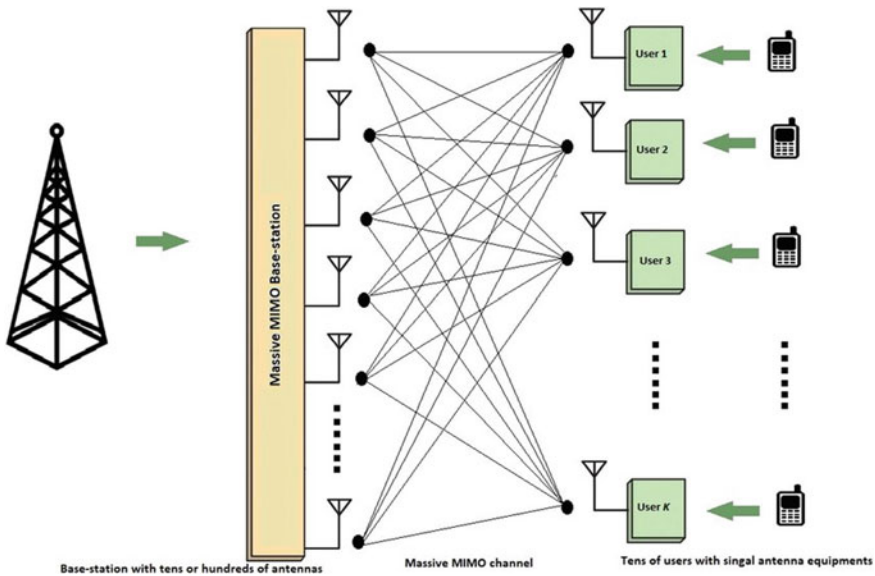
Fig. 5 Illustration of massive MIMO

2.2 Spectrum Sharing

Spectrum sharing is defined as the use of the same frequency band by two or more users (applications) on a nonexclusive basis under a defined sharing arrangement in communication technology [13]. The government of India's Department of Telecommunications (DoT) has also provided information on the rules for spectrum sharing, which allow the use of already available spectrum as well as future frequency bands [14]. In addition to spectrum sharing, the government of India needs to treat the available spectrum with respect to service providers extremely carefully (for the efficient deployment of 5G) as illustrated in Fig. 6.

Spectrum sharing allows users to share a single frequency band with variable priority. The most basic challenge in spectrum sharing is determining when and when it is feasible. This is achieved through spectrum sensing and prediction. Sharing has developed as a practical solution to the spectrum constraint in recent years due to the restricted infrastructure faced by mobile operators due to increased competition. Spectrum sharing has reduced average revenue per user, but capacity demands have increased [15, 16]. Spectrum sharing reduces overall investment and is suited for low traffic areas where network development and maintenance are simple. Smooth multi-operator spectrum sharing benefits dynamic and asymmetric traffic. The most important: obstacle for fairness, data openness, and service quality agreements is building trust between operators [16].
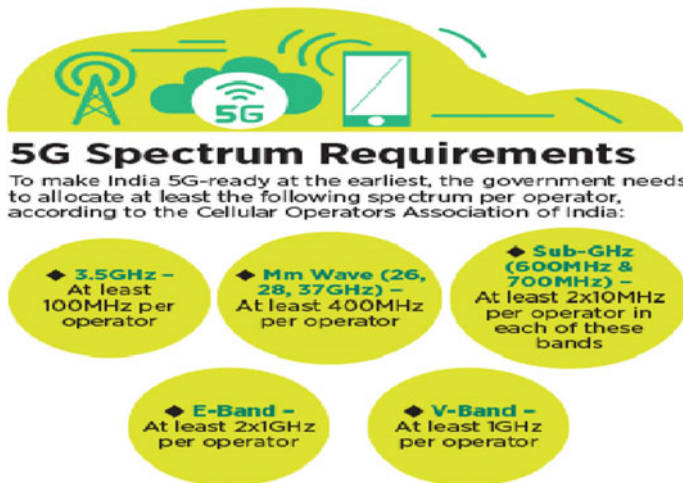
Fig. 6 Necessary divisions for effective spectrum utilization in India

2.3 Interference Management

In the present scenario when work from home becomes an inseparable part of life, wireless traffic increases and hence responsible for Co-Channel Interference (CCI). Whereas densification of network is contemplated as a significant tool to boost traffic capacity and user throughput. Simultaneously when the density and load of the network grow, receiver terminals in the network suffer from high CCI, peculiarly at the boundaries of cells. CCI proves to be the major aspects of averting the further enhancement and development of 4G cellular systems, therefore schemes for interference management are vital [17].

2.4 Advanced Receiver

At the early stage of wireless technology development, the detection of interference was done with respect to thermal noise only. Practically, the interference is detected strongly by the neighboring cells in the communication system. However, the detection of interference via thermal noise weakens the performance capability of the overall system proceedings. Interference randomization techniques were brought into work, but it was profitable only to a moderate level considering quality barriers [18]. The reason being Advance receivers proved to be a better alternative for interference management is least performance loss. Advance receiver proved to be more advantageous because of its characteristic of modulation constellation, symbol detection and decoding as well [17].

2.5 Joint Scheduling

During the releases of LTE versions (8th and 9th), only interference randomization was scrutinized through scrambling of transmitted signals. In the release of 10th and 11th versions (LTE-Advanced), the availability of sufficient headroom for performance enhancement at the edges of the cell was also acknowledged [17].

2.6 Device to Device Communication (D2D)

The cellular users and base stations are interconnected through D2D and that's why D2D technology proves to play an evident role in 5G expansion [19]. The actual work of D2D is to establish interference management in Heterogeneous Network (HetNet). Whereas, HetNet functions as a multi-tier network, which contains high-power nodes. D2D completes the basic requirement for a proper 5G communication
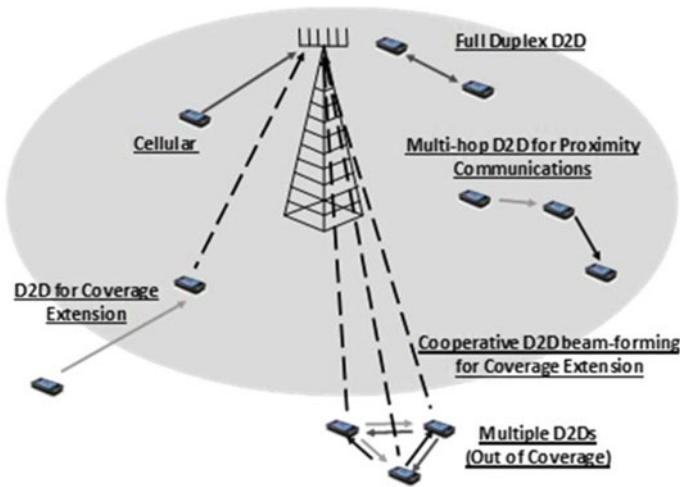
Fig. 7 Extension of coverage

i.e., low end-to-end delay, less power consumption, high spectrum efficiency and high data rates without any system failure but simultaneously compromises the interference management [20–22]. A coverage extension concept using D2D is shown in Fig. 7.

3   Challenges in Implementation of 5G

To bring 5G revolution, the network framework of 4G LTE needs to be secured in the coming time. Technical experts need to present the best possible service as the upcoming technology (5G) has an edge for its promising characteristics and Return on Investments (RoI) [23]. Globally, the challenges faced in implementation of 5G are low latency requirements, building complex and dense network, coverage issues and dealing with new security issues and keeping operating and maintenance cost low. Some challenges from Indian perspective are given below:

Fiber Infrastructure. Fiber infrastructure is one of the significant challenges and its growth will help carrying out 5G more conveniently in India. The significant contribution of fiber infrastructure is to remit increased data range and enhance voice calling capabilities. Because of paucity of fiber cable base in India, people face call drop problems which stipulates the nations low expenditure in fiber and backhaul framework [24, 25]. Only 20% of towers in the country are backhaul in contrast with 80% in countries like China and Korea where they specifically focus on making policy that signifies to fiber categorization [26].

Regulatory Bodies in Telecom Sector. Since the betterment of broadband over last years, in India the telecom sector had overlooked the standard opportunities in mapping out consistent broadband plans. This sector is highly affected by the proceeding delays and their related issuance issues [26].

Low Speed of Data and High Rates. India ranks at 89th position out of 147 countries in terms of average Internet speed of 6.5 mbps (megabits per second) only. Current data speed provided by companies in India is not uniform distinctly in rural areas [26]. The data rate prerequisite for 5G is one Tbps (Terabits per second) so it can be said that India is far behind in terms of average speed and it is mainly because of lack of fiber base.

Availing More Base Stations. More base stations require more antennas and more towers (space and infrastructure) which in turn need more intense radiation and high transmission power. Installing more base station specifically in high population areas is a very tough task as people see it as a danger for living things.

High Consumption of Power. 5G is a technology which consumes comparatively more power than 4G as it provides more network speed and it leads to faster battery drain with respect to usage time. It is a very serious issue for end user's perspective and presently many modifications are going on for the better battery performance.

4 Advanced Technologies Used in 5G

4.1 Internet of Things (IoT)

The International Telecommunication Union's (ITU) vision of IoT is "from anytime, anyplace connectivity for anyone and we will now have the connectivity for anything" [27]. IoT improvises to manage the track inventory more conveniently which lessen ups the human errors. IoT also assists better cost management and optimized resource planning. For example, temperature-monitoring sensors can be utilized for better crops production and send alerts when any unwanted situations occur [28]. In present scenario, IoT implementation (globally) should be capable of providing low latency, high security and complete coverage. In the present situation, 4G is not that efficient to handle this tremendous growth with respect to required Quality of Service (QoS). The sensors used in IoT and its formation should be preferably economic also. As we are coping with IoT systems and it is expected around 80 billion of IoT devices are linked over network [29]. In short, the aim of IoT is to "plug and play smart object" with the following benefits.

- Minimum user interaction
- Proactive maintenance
- New enhanced services
- Utilisation improvement
- Cost reduction.

4.2 Artificial Intelligence (AI)

AI is a technology which helps the users to access the e-commerce activities via supervised marketing decisions and tracking various products simultaneously [30]. The benefits of AI over 5G are expected to be very productive and significant. The main aim of the integration of AI and 5G is to decrease capital expenditure, amplifying network performance and building new revenue streams [31].

4.3 f-OFDM

The term stands for 'Filtered Orthogonal Frequency-Division Multiplexing', which perform division of waveform to numerous sub bands that could be of different bandwidth. Efficiency and protection against interference in fading (simpler channel) equalization are the benefits of f-OFDM. In LTE specifications 10% of system bandwidth is reserved as guard bands to minimize the adjacent channel leakage ratio, avoiding over lapping of sub bands and fulfill spectrum mask requirements. By applying f-OFDM the baseband OFDM signal can be shaped with ultra-narrow sub region and can meet the requirement of LTE spectrum along with other bandwidth related requirements [32, 33].

4.4 Virtual Reality (VR) and Augmented Reality (AR)

VR and AR come out to be the driving forces in the commercial market globally [34]. Presently the standards of 4G abide from restrictions like latency, speed and bandwidth and considering these concerns, 5G is almost ready to unlatch the full proficiency of VR and AR technologies. These weaknesses can be reduced by enhancing network speed and lower latency. According to time trade research 85% of users go for shopping in physical stores and many features of physical shopping cannot be exchanged by e-commerce but AR and VR can create user experiences that are much closer to physical shopping. VR and AR with 5G have the capability to bring the huge change in the e-commerce activities in the upcoming years subject to smooth and efficient deployment of 5G [35]. A brief description is given in Table 1.

5 Security Issues in 5G

5G systems must adopt security features to enhance level of trust among subscribers [36]. Some of the important aspects are discussed below:

Table 1 Usage of recent technologies in 5G

S. No.	Technology	Impact	Advantages of 5G
1	(IoT)	Upgrade user acquaintance, manage orders more productively	Transfer of data by IoT devices will be made more favorable
2	Artificial intelligence (AI)	Performing e-commerce activities and order products online	Easy and quick access of information to acknowledge the particular context in a better way
3	f-OFDM	Enable spectrum slicing and coexistence of multiple sub bands	More substantial to frequency selective fading comparatively to single carrier systems
4	Virtual and augmented reality	Enhanced interaction with reality	Make physical shopping easier and improvise latency requirements

Inter Operator Security: In previous 2G/3G and 4G networks, some critical loopholes were found based on SS7 architecture and diameter protocols etc. 5G must adopt inter operator security along with the inclusion of more trustworthy proxy servers.

Key and Authentication: It will be better if 5G systems always use the public key of available home network for the required encryption. The entire network and devices must be mutually authenticated and the data transmission beyond the domain of current service provider or in case of wi-fi calling, the feature of secondary authentication must be added.

5G and IoT: The IoT devices using 5G consists of various hardware and software and it leads to security problems because many of them may be vulnerable to Man in the Middle Attack (MITM), replay attacks and password guessing attacks etc. Some of these devices will also transfer the personal information of user such as name, Date of Birth (DoB), contact number of user and credit card details etc. which in turn creates security vulnerabilities. Table 2 describes possible attacks in 5G with brief description [37].

6 Conclusion and Future Scope

A comprehensive review with respect to evolution of 5G and associated security issues, emerging technologies and impact is done in this paper. The expectations from 5G, market opportunities and expected growth are discussed. It is also very interesting to see that various service providers have different plans related to 5G services. The Indian 5G market is expected to grow 11% by 2025. The use of massive MIMO, spectrum sharing and interference management will be the key factors. Issues

Table 2 Possible security attacks with reason

S. No.	Attack type	Reason
1	Eavesdropping	Intruder can eavesdrop the transferred message and it will be helpful for launching other attacks
2	Traffic analysis	Intruder can observe the ongoing traffic in order to predict other useful parameters
3	Replay attack	Intruder can intercept the ongoing communication and can perform delay or deceitful retransmission
4	MITM	Intruder can be there between the ongoing transmission of two devices and modifies the messages [38]
5	Impersonation attack	Intruder can find the identity of any legitimate participating entity and send malicious messages on its behalf
6	Denial of service (DoS) attack	Intruder can prevent the legitimate entities from using the available resources [39]
7	Database attack	In 5G-IoT based systems, intruder can enjoy the fact that the database is managed through various servers and due to this, so many attacks in this category are possible
8	Malware attack	Intruder can run malicious program in a targeted system so that many unauthorized activities can be performed
9	Insider attack	An insider means existing user of a trusted system can use the available information maliciously and it can lead to other potential attacks also
10	Physical access	An intruder can access to an IoT device physically and some sensitive information (like passwords, session keys etc.) can be accessed directly

related to D2D communication needs to be planned carefully. The growth of fiber infrastructure and regulations imposed by Telecom Regulatory Authority of India (TRAI) will be the deciding factors in India. The growth of IoT is also dependent on 5G and the use of VR and AR will be even more interesting in the overall deployment of 5G. As long as security issues of 5G are concerned, there are some serious risks.

Inter operator security, key management along with security attacks are some of the very important points. Security attacks include a wide variety like MITM, DoS or physical access to IoT devices, etc. The success of 5G will be based on the fact that how these security issues are addressed keeping the overall complexity as minimum as possible. Needless to mention that the future of 5G will be very promising and it has the potential to develop new jobs and market opportunities. The amalgam of 5G and IoT will provide new level of user experience and it will definitely benefit the entire society at large. We also invite research fraternity to throw more light on 5G and associated technologies with related security issues as it is going to hit the market very soon and it can be the fruitful result of this paper.

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Oughton EJ, Lehr W, Katsaros K, Selinis I, Bubley D, Kusuma J (2021) Revisiting wireless internet connectivity: 5G vs wi-fi 6. Telecommun Policy 45(5):1–15
2. Fang H, Wang X, Tomasin S (2019) Machine learning for intelligent authentication in 5G and beyond wireless networks. IEEE Wirel Commun 26(5):55–61
3. Temesvári ZM, Maros D, Kádára P (2019) Review of mobile communication and the 5G in manufacturing. Procedia Manuf 32:600–612
4. Nagul S (2018) A review on 5G modulation schemes and their comparisons for future wireless communications. In: Conference on signal processing and communication engineering systems, pp 72–76
5. Gupta A, Raghav K, Dhakad P (2019) The effect on the telecom industry and consumers after the introduction of reliance jio. Int J Eng Manag Res 9(3):118–137
6. Puri S, Rai RS, Saxena K (2018) Barricades in network transformation from 4G to 5G in India. In: 7th international conference on reliability, infocom technologies and optimization (trends and future directions), pp 695–702
7. Ericsson Mobility Report: 5G subscriptions to top 2.6 billion by end of 2025. Available at: https://www.ericsson.com/en/press-releases/2019/11/ericsson-mobility-report-5g-subscriptions-to-top-2.6-billion-by-end-of-2025
8. Shafique K, Khawaja BA, Sabir F, Qazi S, Mustaqim M (2020) Internet of Things (IoT) for next-generation smart systems: a review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. IEEE Access 8:23022–23040
9. Wang J, Wang G, Li B, Yang H, Hu Y, Schmeink A (2021) Massive MIMO two-way relaying systems with SWIPT in IoT networks. IEEE Internet Thing J 8(20):15126–15139
10. Gupta A, Jha RK (2015) A survey of 5G network: architecture and emerging technologies. IEEE Access 3:1206–1232
11. Larsson EG, Edfors O, Tufvesson F, Marzetta TL (2014) Massive MIMO for next generation wireless systems. IEEE Commun Mag 52(2):186–195
12. Chataut R, Akl R (2020) Massive MIMO systems for 5G and beyond networks—overview, recent trends, challenges, and future research direction. Sensors 20(10):1–35
13. Kibria MG, Villardi GP, Ishizu K, Kojima F, Yano H (2016) Resource allocation in shared spectrum access communications for operators with diverse service requirements. EURASIP J Adv Signal Process 2016(Article Number 83):1–15
14. Spectrum sharing guidelines 2021. Department of telecommunications, Ministry of communications, Government of India. Available at: https://dot.gov.in/spectrummanagement/spectrum-sharing-guidelines-2021
15. Ahmed F, Kliks A, Goratti L, Khan SN (2018) Towards spectrum sharing in virtualized networks: a survey and an outlook, cognitive radio, mobile communications and wireless networks (part of the EAI/Springer innovations in communication and computing book series), pp 1–28
16. Customer data: designing for transparency and trust, analytics and data science. Harvard Bus Rev. Available at: https://hbr.org/2015/05/customer-data-designing-for-transparency-and-trust
17. Nam W, Bai D, Lee J, Kang I (2014) Advanced interference management for 5G cellular networks. IEEE Commun Mag 52(5):52–60
18. Xu Z, Xu G, Zheng Z (2021) BER and channel capacity performance of an FSO communication system over atmospheric turbulence with different types of noise. Sensors 21(10):1–14
19. Adnan MH, Zukarnain ZA (2020) Device-To-Device communication in 5G environment: issues, solutions, and challenges. Symmetry 12(11):1–22

20. Alzoubi KH, Roslee MB, Elgamati MAA (2019) Interference management of D2D communication in 5G cellular network. In: Symposium on future telecommunication technologies, pp 1–7
21. Xu Y, Liu F, Wu P (2018) Interference management for D2D communications in heterogeneous cellular networks. Pervasive Mob Comput 51:138–149
22. Wei L, Hu RQ, Qian Y, Wu G (2014) Enable device-to-device communications underlaying cellular networks: challenges and research aspects. IEEE Commun Mag 52(6):90–96
23. Westberg E, Staudinger J, Annes J, Shilimkar V (2019) 5G infrastructure RF solutions: challenges and opportunities. IEEE Microwave Mag 20(12):51–58
24. Fiber investments key to success of 5G in India. Available at: https://telecom.economictimes.indiatimes.com/tele-talk/fibre-investments-key-to-success-of-5g-in-india/2452
25. Boateng ON, Xedagbui FEB, Adekoya AF, Weyori BA (2020) Fiber optic deployment challenges and their management in a developing country: a tutorial and case study in Ghana. Eng Rep 2(2):1–16
26. Sharma S (2018) Problems in implementing 5G in India and solutions for it. Int J Manag Appl Sci 4(5):78–82. Available at: http://www.iraj.in/journal/journal_file/journal_pdf/14-469-153303674878-82.pdf
27. Madakam S, Ramaswamy R, Tripathi S (2015) Internet of Things (IoT): a literature review. J Comput Commun 3(5):164–173
28. Kshetri N (2018) 5G in e-commerce activities. IT professional 20(4):73–77
29. Chettri L, Bera R (2020) A comprehensive survey on internet of things (IoT) toward 5G wireless systems. IEEE Internet Things J 7(1):16–32
30. Davenport T, Guha A, Grewal D, Bressgott T (2020) How artificial intelligence will change the future of marketing. J Acad Mark Sci 48:24–42
31. Haidine A, Salmam FZ, Aqqal A, Dahbi A (2020) Artificial intelligence and machine learning in 5G and beyond: a survey and perspectives. Moving broadband mobile communications forward—intelligent technologies for 5G and beyond, pp 1–21
32. Taher MA, Radhi HS, Jameil AK (2021) Enhanced F-OFDM candidate for 5G applications. J Ambient Intell Hum Comput 12:635–652
33. Figueiredo FAPD, Aniceto NFT, Seki J, Moerman I, Fraidenraich G (2019) Comparing f-OFDM and OFDM performance for MIMO systems considering a 5G scenario. In: IEEE 2nd 5G world forum, pp 532–535
34. Flavián C, Sánchez SL, Orús C (2019) The impact of virtual, augmented and mixed reality technologies on the customer experience. J Bus Res 100:547–560
35. Attaran M (2021) The impact of 5G on the evolution of intelligent automation and industry digitization. J Ambient Intell Hum Comput 1–17.
36. Khan R, Kumar P, Jayakody DNK, Liyanage M (2020) A survey on security and privacy of 5G technologies: potential solutions, recent advancements, and future directions. IEEE Commun Surv Tutor 22(1):196–248
37. Wazid M, Das AK, Shetty S, Gope P, Rodrigues JJPC (2020) Security in 5G-enabled IoT communication: issues, challenges, and future research roadmap. IEEE Access 9:4466–4489
38. Shukla V, Chaturvedi A, Srivastava N (2019) A secure stop and wait communication protocol for disturbed networks. Wirel Pers Commun 110:861–872
39. Shukla V, Mishra A (2020) A new sequential coding method for secure data communication. In: IEEE international conference on computing, power and communication technologies, pp 529–533

A Secure Multi-tier Framework for Healthcare Application Using Blockchain and IPFS

H. M. Ramalingam, H. R. Nagesh, and M. Pallikonda Rajasekaran

Abstract Information technology advancement and the requirements in health care are to provide a variety of solutions to solve different problems. The better utilization of the technological advances of cloud computing, sensor networks, distributed processing, and distributed storage, security, and privacy in healthcare applications will give real-time monitoring of the patient's vital signs, storage of healthcare data, and safe and continuous access of health records. This paper implemented a multi-tier framework that comprises the best features like blockchain-based security of health records, JPPF-based parallel processing, and decentralized storage using IPFS. Also, this framework provides the facility for remote patient monitoring.

Keywords Electronic health records · Blockchain · IPFS

1 Introduction

In the last few years, the growth of wireless sensor networks, wearable IoT devices has enhanced the eminence patient care through remote patient monitoring [1]. IoT consists of smart devices capable of sensing and transmitting the data to the cloud via IoT Gateways, which can be analyzed in the cloud or analyzed locally while creating a health monitoring framework. The drawbacks of the IoT need to be addressed. As the number of interconnected devices increases, information sharing will be challenging in storage, processing, and security. Also, there is no derived international standard for IoT. So, it's difficult for making different devices to communicate flawlessly. The primary requirements for a better healthcare framework need to address the

H. M. Ramalingam (✉)
Mangalore Institute of Technology and Engineering, Moodabidri, India
e-mail: ramalingam@mite.ac.in

H. R. Nagesh
Sahyadri College of Engineering & Management, Mangalore, India

M. P. Rajasekaran
Kalasalingam Academy of Research and Education, Krishnankoil, India

continuous monitoring of the patient's health status, storage processing of the data, privacy, and security of the patient's health records.

A relationship between wearable sensors and distributed computing will be an excellent area to discover. The increased number of wearable sensors in health care transmits a massive amount of data, which challenges storage, processing, and security in a real-time environment. These problems can be fixed by introducing distributed scenario for data processing and a decentralized plan for storage [2]. The wearable sensors would collect vast numbers of data processed and stored efficiently in high-performance and cloud-based systems by combining these systems. The mixture of such a system, the wireless sensor cloud, would deliver a powerful platform for research studies, medical applications, extensive military operations, and commercial establishments [3]. Implementing this distributed processing will also enable the computing needs for the analysis software.

Health record management is one more challenge in healthcare applications regarding security and privacy. This problem can be overcome by improving the health record management scheme using blockchain [4, 5]. Blockchain stores the information in a distributed and immutable manner, giving enhanced security and privacy. Since the blockchain is immutable, removing health records from the system is complex, so to make the system more viable, we have to consider an alternate method of storing health records without losing the benefits of blockchain. The most pleasing way to do this is to use an interplanetary file system (IPFS) to keep the heath records and a blockchain to store the index of the file and the user information.

By considering the requirements of the best healthcare application, we implemented a multi-tier framework that comprises the best features like blockchain-based security of health records, JPPF-based parallel processing, and decentralized storage using IPFS. Also, this framework provides the facility for remote patient monitoring.

2 Related Works

A remote patient monitoring system with blockchain-based security and privacy solution for the IoT is implemented in [1]; here, the author presented a hybrid approach consisting of a private key, a public key, and a blockchain to ensure patient-centric access control. A comprehensive study on health care 4.0 by Hathaliya and Sudeep [2]. Given the healthcare industry's current requirements regarding security and privacy, they have also explained various existing methods and their merits and demerits.

Security and privacy requirements in the future e-health direction of the healthcare sector are explained in [6, 7]. Security comparison and performance analysis of different methods for IoT data storage are presented in [8–11]. The authors in [12–14] demonstrated the working of smart contracts and implemented the blockchain-based medical record sharing system, which provides security and privacy for patient details. The usage of IPFS and blockchain and their advantages are showcased in [15, 16].

The usage of wireless wearable devices in remote patient monitoring and the authors [17] proposed a three-tier architecture for early heart defect diagnosis. The sensor grid utilization in patient monitoring and the usage of grid computing to enhance the performance of wireless sensor information collection and analysis are explained in [18]. An extensive overview of sensor cloud architecture, implementation of the same in health care, and comparison of diverse methods in sensor-cloud were reviewed by the authors in [19]. Connecting various IoT devices and its standard and protocols using the existing web technologies is demonstrated by the authors in [20], and in [21] the authors showcased a security framework for health applications, which gives the clear idea of creating a specialized framework with multiple layers to improvise the system requirements for healthcare applications. The authors implemented architecture to support healthcare applications using JPPF-based parallel processing [3, 22].

This extensive study about the related work gives an idea to use the best features of these prosed methods to create a new framework to serve the healthcare applications, volume editors. Usually, the program chairs will be your main points of contact to prepare the volume.

3 Methodology

The critical feature of the proposed model is to use the available resources, which makes any system a master driver to distribute the tasks, and the connected system in the network acts as the worker nodes to complete the assigned tasks. Also, a private IPFS is configured within the local networked systems to store the health records in a distributed manner. Figure 1 represents the proposed multi-tier architecture. We are creating similar structures in different hospitals and creating a community cloud to share the services, provide on-demand healthcare services, and remote patient monitoring.

The Java Parallel Processing Framework (JPPF) usage will help reduce the processing time for any analysis software and efficiently handle the continuous monitoring data received from the wearable devices. JPPF splits any provided application into smaller parts and runs simultaneously on various machines to reduce processing time, which helps to make easy processing of real-time sensory data. The JPPF architecture and the work distribution process are represented in Fig. 2.

The blockchain-based register is kept to guarantee the privacy and security of the patient information and electronic health record, which has all the details of patient and doctor. One more major critical element of the architecture is the introduction of the IPFS for distributed file storage. When so ever we are concerned about the privacy of the patient's health records, the system should provide an option to delete and update the health records wherever necessary. Since the blockchain is immutable, it is technically impossible to erase or modify the health records if we directly store them in the blockchain network. Our architecture provides one more tier of distributed storage using IPFS. Figure 3 depicts the typical workflow of the platform. In this
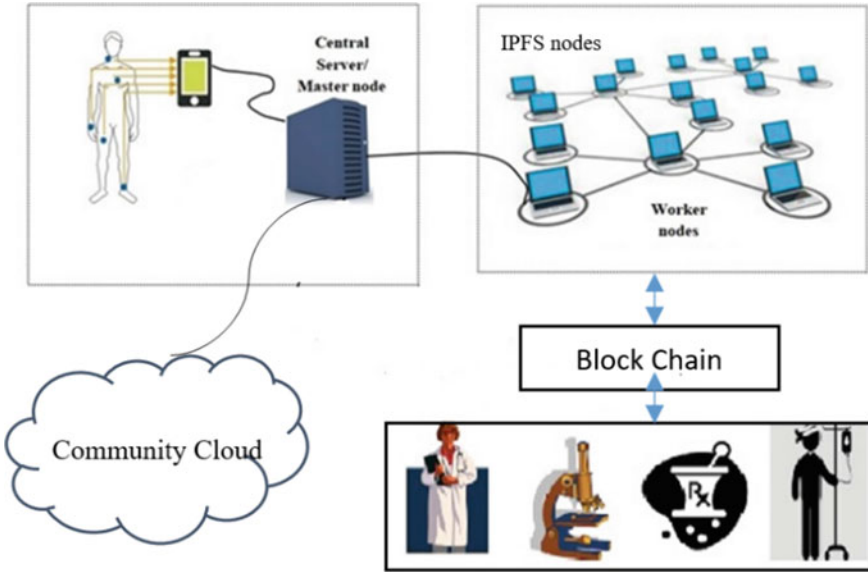
Fig. 1 Proposed multi-tier architecture
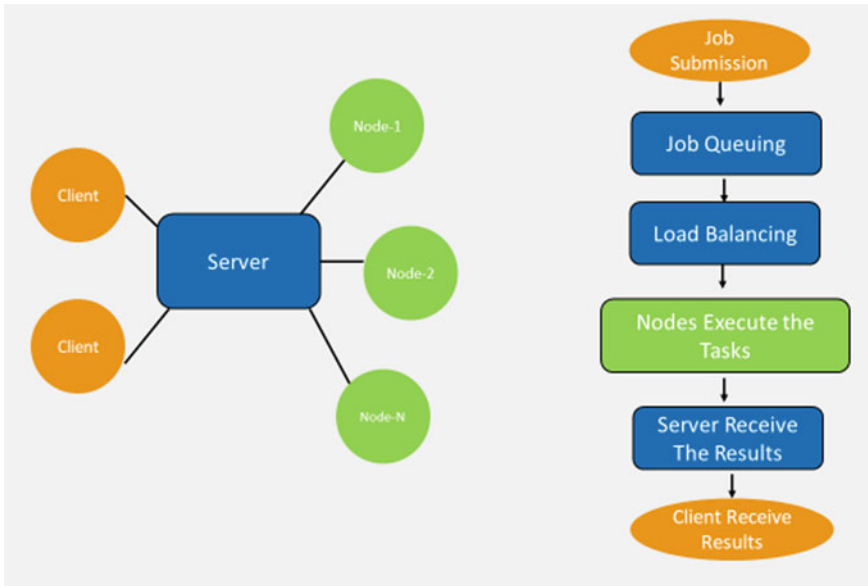
Fig. 2 Architecture and the work distribution on JPPF
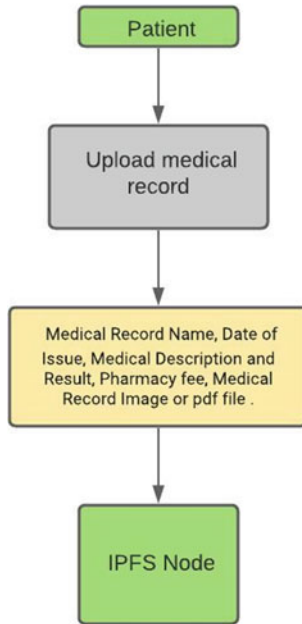
Fig. 3 Workflow of the blockchain and IPFS-based platform to add health records

storage method, the health records added by the patient will be encrypted by the patient Id and stored in the IPFS network, which provides an SHA-256 hash as the index of the file. The hash value will be held in the blockchain smart contract.

4 Results and Discussion

The experimental setup has been implemented using systems having intel dual-core processors running with Ubuntu Linux in a 2 GB RAM configuration, where the JPPF master and slave nodes are installed. These systems are LAN connected by a fast ethernet switch. Figure 4 depicts the implementation monitoring in the JPPF driver node. Also, for the distributed storage, IPFS is configured among the systems.

The parallel processing scheme has been evaluated by running a 300×300 matrix multiplication of 100 iterations within the JPPF configured systems with a dynamic task allocation scheme. Figure 5 shows the performance comparison of the same, the efficiency of the parallel processing depends on the availability of the idle resources and the network condition. The number of worker nodes increased the execution time drastically reduced. This implementation helps any analysis software run in a distributed manner and handle real-time senor data efficiently.

React.js and web3.js have created the health record maintenance user interface and the live sensory data visualization, and the meta mask extension is also used
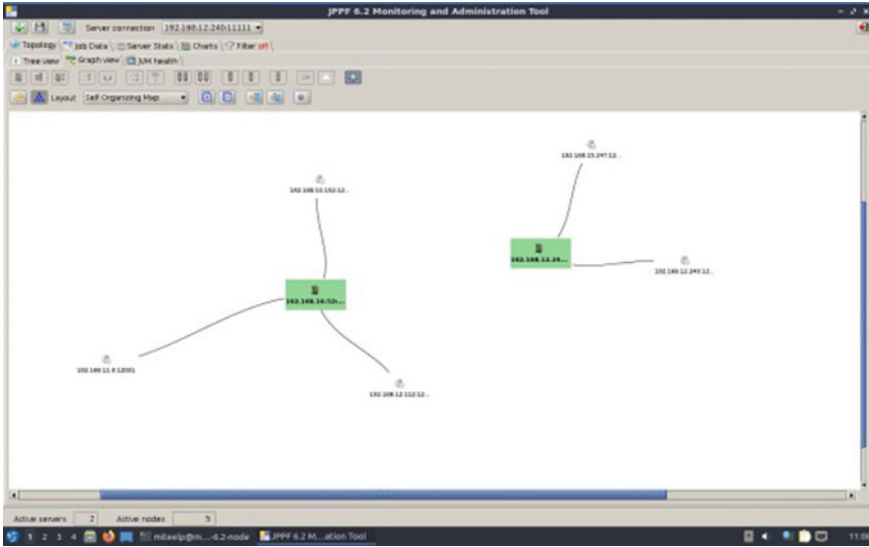
Fig. 4 JPPF driver and nodes are running in different systems with local IPFS configured
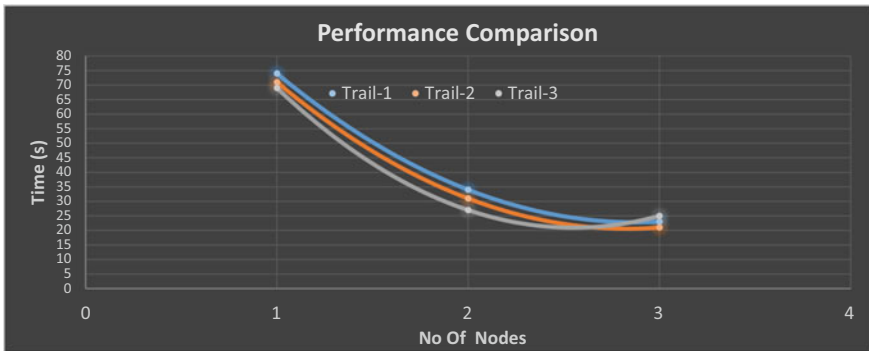
Fig. 5 Parallel processing performance comparison

for the web interface to interact with the Ethereum blockchain. The user interface is depicted in Fig. 6. In order to mimic the blockchain network, ganache-cli is used. The use of blockchain makes sure the security and privacy of the patients. The health records are stored and distributed using the IPFS configuration among the systems.

The significant advantage of using distributed health record storage is improved security and increased data access speed. The IPFS storage mechanism gave a much faster retrieval time than other file transfer methods. Figure 7 presents the data retrieval comparison in the same network environment with different accessing schemes. For this, a Raw image of the size of around 50 MB is used.
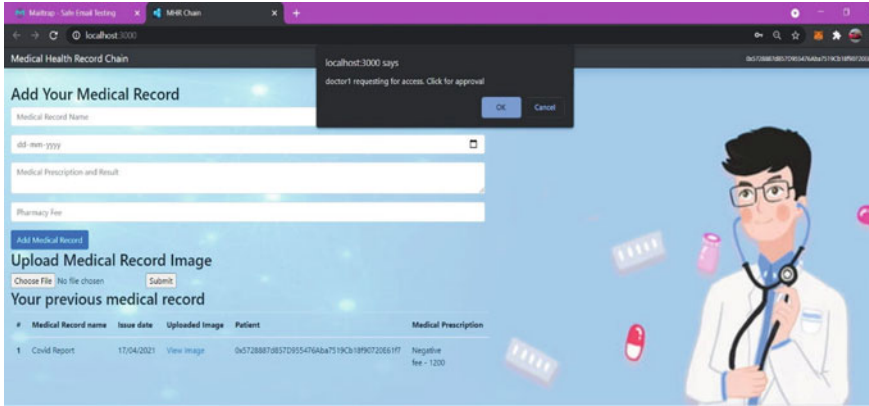
Fig. 6 User interface to share health records and IoT data
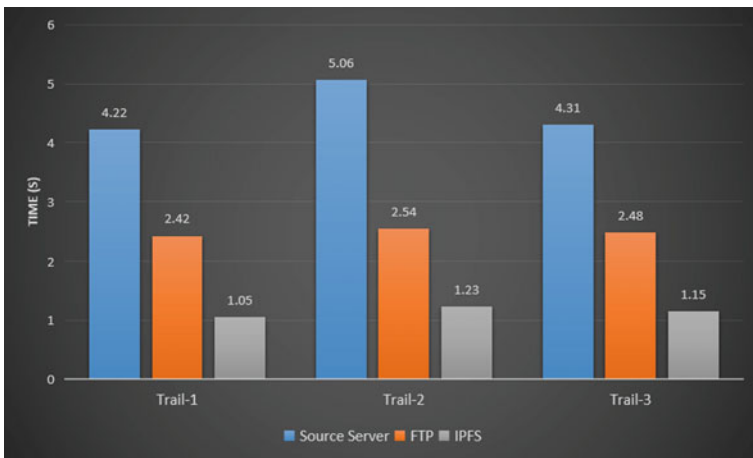
Fig. 7 Data retrieval time comparison

5 Conclusion

We have proposed a multi-tier architecture with various technologies for healthcare applications to enhance the privacy and security of medical information. Our experimental study demonstrated the architecture of blockchain-based security for health records, JPPF-based parallel processing, and decentralized storage using IPFS. Also, this framework provides the facility for remote patient monitoring with a secured IoT data transfer.

References

1. Dwivedi AD, Srivastava G, Dhar S, Singh R (2019) A decentralized privacy-preserving healthcare blockchain for IoT Sens 19:1–17. https://doi.org/10.3390/s19020326
2. Hathaliya JJ, Tanwar S (2020) An exhaustive survey on security and privacy issues in healthcare 4.0. Comput Commun 153:311–335. https://doi.org/10.1016/j.comcom.2020.02.018
3. Ramalingam HM, Rajasekaran MP, Nagesh HR (2020) WBAN implementation in a parallel processing environment for E-healthcare applications. Int J Future Gen Commun Network 13:4963–4969
4. Tanwar S, Parekh K, Evans R (2020) Blockchain-based electronic healthcare record system for healthcare 4.0 applications. J Inf Secur Appl 50. https://doi.org/10.1016/j.jisa.2019.102407
5. Sharma Y, Balamurugan B (2020) Preserving the privacy of electronic health records using blockchain. Proc Comput Sci 173:171–180. https://doi.org/10.1016/j.procs.2020.06.021
6. Azeez NA, van der Vyver C (2019) Security and privacy issues in e-health cloud-based system: a comprehensive content analysis. Egypt Inf J 20:97–108. https://doi.org/10.1016/j.eij.2018.12.001
7. Hussien HM, Yasin SM, Udzir SNI et al (2019) A systematic review for enabling of develop a blockchain technology in healthcare application: taxonomy, substantially analysis, motivations, challenges, recommendations and future direction. J Med Syst 43. https://doi.org/10.1007/s10916-019-1445-8
8. Liu YH, Zhang S (2020) Information security and storage of Internet of things based on blockchains. Future Gen Comput Syst 106:296–303. https://doi.org/10.1016/j.future.2020.01.023
9. Roehrs A, da Costa CA, da Rosa RR et al (2019) Analyzing the performance of a blockchain-based personal health record implementation. J Biomed Inform 92:103140. https://doi.org/10.1016/j.jbi.2019.103140
10. Tang F, Ma S, Xiang Y, Lin C (2019) An efficient authentication scheme for blockchain based electronic health records. IEEE Access 7:41678–41689. https://doi.org/10.1109/ACCESS.2019.2904300
11. Deebak BD, Al-Turjman F, Aloqaily M, Alfandi O (2019) An authentic-based privacy preservation protocol for smart e-healthcare systems in IoT. IEEE Access 7:135632–135649. https://doi.org/10.1109/ACCESS.2019.2941575
12. Sun J, Ren L, Wang S, Yao X (2020) A blockchain-based framework for electronic medical records sharing with fine-grained access control. PLoS ONE 15:1–23. https://doi.org/10.1371/journal.pone.0239946
13. Shahnaz. A Qamar. U, Khalid.A (2019) Using blockchain for electronic health records. IEEE Access 7:147782–147795. https://doi.org/10.1109/ACCESS.2019.2946373
14. Chakraborty S, Aich S, Kim HC (2019) A secure healthcare system design framework using blockchain technology. In: International conference on advanced communication technology, ICACT 2019, pp 260–264. https://doi.org/10.23919/ICACT.2019.8701983
15. Miyachi K, Mackey TK (2021) hOCBS: A privacy-preserving blockchain framework for healthcare data leveraging an on-chain and off-chain system design. Inf Process Manage 58:102535. https://doi.org/10.1016/j.ipm.2021.102535
16. Reen GS, Mohandas M, Venkatesan S (2019) "Decentralized patient-centric e-Health record management system using blockchain and IPFS. In: 2019-IEEE conference on information and communication technology, CICT 2019. https://doi.org/10.1109/CICT48419.2019.9066212
17. Sonal, Reddy SRN, Kumar D (2020) Early congenital heart defect diagnosis in neonates using novel WBAN based three tier network architecture. J King-Saud Univ Comput Inf Sci 1–12. https://doi.org/10.1016/j.jksuci.2020.07.001
18. Pallikonda Rajasekaran M, Radhakrishnan S, Subbaraj P (2010) Sensor grid applications in patient monitoring. Future Gen Comput Syst 26:569–575. https://doi.org/10.1016/j.future.2009.11.001
19. Alamri A, Ansari WS, Hassan MM et al (2013) A survey on sensor-cloud: architecture, applications, and approaches. Int J Distr Sens Netw

20. Pasha M, Shah SMW (2018) Framework for E-health systems in IoT-based environments. Wirel Commun Mob Comput. https://doi.org/10.1155/2018/6183732
21. Hussain A, Ali T, Althobiani F et al (2021) Security framework for IoT based real-time health applications. Electronics (Switz) 10:1–15. https://doi.org/10.3390/electronics10060719
22. Ramalingam HM, Nagesh HR, Pallikonda Rajasekaran M (2020) Experimental study of high-performance computing in three-tier architecture for E-health care application. Int J Adv Sci Technol 29:11078–11085

Machine Learning Techniques to Web-Based Intelligent Learning Diagnosis System

Ch. Ravisheker, M. Prabhakar, Hareram Singh, and Y. Shyam Sundar

Abstract In this paper, machine learning techniques to web-based intelligent learning diagnosis system are implemented. The main intent of this paper is to cultivate the ability of learner's knowledge. This is done by integrating number of opportunities to the learner. This method helps the learner to improve knowledge and ability to work on diagnosis system. Diagnosis system will predict the results in very effective way. Initially, input data is mapped based on features using data sample mapping procedure. Next, the mapped data is classified using optimal classification technique. The optimal classification technique is based on the features. This classification techniques is performs its operation in two ways they are testing and training. The tested and trained data extracts its features using feature extracted method. Both SVM and ELM will classify the data based on extraction. At last, data is evaluated and classified. From results, it can observe that accuracy, reliability, precision, recall, $F1$ score, and mean gives effective outcome compared with Naïve Bayesian classifier.

Keywords Web-based learning · Theme-based learning · Naïve Bayesian classifier · Support vector machines · Learning diagnosis

Ch. Ravisheker (✉) · M. Prabhakar · Y. Shyam Sundar
CMR Engineering College, Hyderabad, Telangana, India
e-mail: ravisheker.ch@cmrec.ac.in

M. Prabhakar
e-mail: prabhakar.m@cmrec.ac.in

Y. Shyam Sundar
e-mail: shyamsundar.y@cmrec.ac.in

H. Singh
GNIOT, Greater Noida, India

717

1 Introduction

In the network learning, a new vision is developed based on the information technology. This learning will help the learner to facilitate the education in effective way. For the generation of useful knowledge which is competent of separately attaining data as well assimilating that data is said to be the process of machine learning. The implementation as well as idea of machine learning is done by calculating the software organization usually acts as a human being that is learned from experience, also consider the interpretations prepared and reforms affording the improved effectiveness and efficiency [1].

The procedures by which nearly all up to date models are developed are adaptive, based upon the analysis of statistics and model performance, numerical methods are used to adjust the model parameters. Through the scrutiny of data and introspection, machine learning procedures create an analytical representation without human involvement model parameters are determined. The examination of how the programs can be learned and written is defined by the machine learning which is said to be the area of Artificial Intelligence (AI). Prediction or classification is often used by the machine learning in the data mining. The processor composes a forecast with machine learning and then based on the criticism as if it is right "attains" from this response. Throughout illustrations of feedback and concern knowledge, it will be learned.

While a comparable circumstance occur in the prospect, this response is useful to make a completely different prediction or to make the same prediction. The important part in machine learning programs is the statistics as the outcome of the forecast must be pretentiously important and have to execute better than an inexperienced forecast. The typical use of machine learning techniques in which the applications include training moving robots, speech recognition, game playing, and classification of astronomical structures. For the representation of the data, a model is used (like a coherent formation similar to a decision tree or a neural network), while data mining tasks are applied by machine learning.

A sample of the database is used to perform the preferred task during the learning process to train the system properly. Next, the general database is activated by the system to truly execute the task. The division of this extrapolative modeling access is taken into two phases. A model that is created for sampled data or historical data is used during the training phase represents the data. It is general that this model is classical data for the database as a whole but also for this sample data and for future data as well. This model is then applied by the testing phase to the left over and outlook data.

Learning process which is stimulated by the human brain with an artificial representation is called an artificial network. Artificial neural network (ANN) is otherwise said to be as neural network [2]. Resembling efficient information with the biological neural network along with its characteristics is known as artificial neural network. The characterization of the collective behavior is by their ability to recall, learn as

well as similar data to that of the human brain or the pattern which is generalized for the training.

Neural networks are also referred as artificial neural networks which are in non-predictive mode and these are being modeled after the functioning of the human brain, which are differentiated from the biological neural networks. There are various algorithms and a graph that represents the processing system that access the graph which is all done with an information processing system in the neural networks. The structure of the neural networks is the directed graph with various intersections (processing elements) as well as arcs (interconnections) between them, which consist of various linked processing elements as with the human brain.

The main use of the neural networks which also uses local data (input and output to the node) [3] is facilitated by this feature in a parallel and/or distributed environment in which the function of the processing elements is independent from the other to direct its processing. However, the arcs in the graph are interconnections and nodes are like individual neurons.

The approach of neural networks to build the representing model will be decision trees which a graphical structure is required and then this structure is applied to the data [4]. The view of the neural networks is in three ways which is as a directed graph with sink (output), source (input), and internal (hidden) nodes. In input layer, the nodes of the input exist, while in the output layer, nodes of the output exist. Over one or more hidden layers the hidden nodes exist. The network of the typical feed forward which has neurons arranged in a distinct layered topology. Layer of the input is not really neural at all: the values of the input variables are introduced by these units with simple serve. In the preceding layer, the neurons of the hidden and output layers are each linked to all of the units [5]. Once more, to determine networks it is possible that they are only few units that are partially connected in the preceding layer yet, for nearly each of the applications fully connected networks are better.

2 Machine Learning Algorithms

2.1 Artificial Neural Network (ANN)

It is an exclusive intelligence tool, which is used for simulation of human brain to generate results and for analysis. This ANN algorithm is a multilayer neural network, has excellent features in non-linear mapping, self-organization, and generalization. This algorithm gives reasonable solutions when presented with inputs, which was un-seen in neural networks [6]. To train a network on a representative set of input–output pairs is possible by the generalization property. Training the network for all input–output pairs can obtain good results.

2.2   Decision Tree (DT)

A desired effect that can be produced by the predicting future outcomes and identifying are usually the most important intentions of data mining and data analysis. One of the most popular predictive modeling methods is the decision trees for the purposes of data mining because interpretable rules and logic statements are provided by them that enable more decision making which is intelligent [7]. The representation of a predictive model which is a tree shaped as well used in the accumulation, classification and forecast tasks is known as decision tree. A technique that is used by the decision tree is a "divide and conquer" that splits the crisis explore gap into subsets. A decision tree is a predictive model in machine learning and data mining that maps conclusions about its target value from the observations about an item.

For such tree models, the more descriptive names are the classification tree. Whereas the leaves of the tree represent the division of the confidential material, then every bough of the tree enacts a sorting query in a decision tree [8]. The tasks that are related to clustering and classification are generally suitable by the decision trees. To explain the decision being taken it helps create rules which are used. The cast between several courses of action, decision trees are excellent tools for helping. A highly efficient framework is provided in which to present options and examine the probable reactions of deciding these preferences. Helping is done to create a reasonable portrait of confronts and opportunities allied with every probable way of achievement.

2.3   Naïve Bayes (NB)

The main intent of Naïve Bayes is to classify the text which is taken from dataset. In classification algorithm, Naïve Bayes is supervised machine learning technique. Examples of Naïve Bayes are Spam flit ration, sentimental analysis, etc. Naïve Bayes algorithm depends upon the attributes of independent variables.

2.4   Random Forest (RF)

For the purpose of both classification and regression in supervised machine learning random forest is utilized. Forest is nothing but group of trees in random forest algorithm [9]. The main intent of random forest is to classify the data. Here, trees are called as decision tress. Accurate output will be obtained when decision tress are higher in number. In random forest algorithm, initially random samples are collected and then decision tress will be created based on existing trees [10].

3 Web-Based Intelligent Learning Diagnosis System

Figure 1 shows the flow chart of web-based intelligent diagnosis system.

The entire flow chart performs its operation in three phases. In first phase, the dataset is investigated based on module and captures the data from dataset.

In next phase, the quality of data is enhanced, and at last, in third phase the noise is deleted in the data.

Algorithm

STEP-1 Initially, input data is mapped based on features using data sample mapping procedure.

STEP-2 Next the mapped data is classified using optimal classification technique.

STEP-3 The optimal classification technique is based on the features.

Fig. 1 Flowchart of web-based intelligent diagnosis system
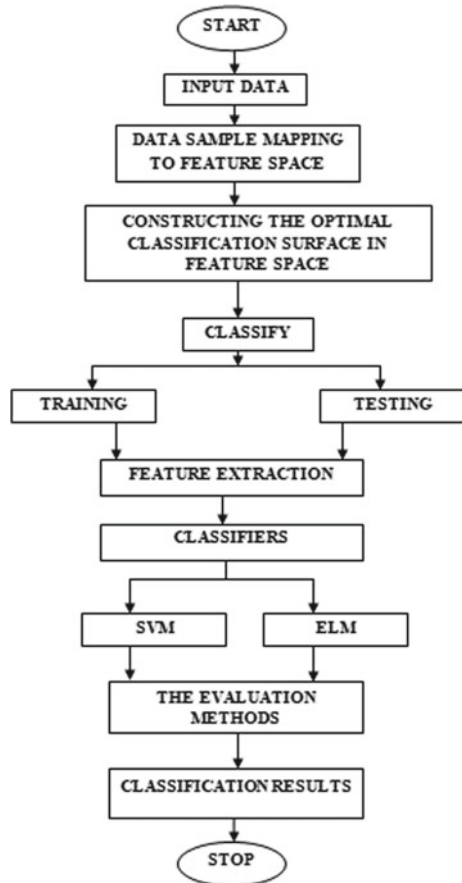
Fig. 2 Example of SVM
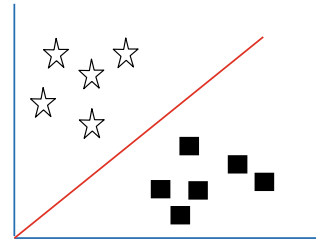
STEP-4 This classification technique performs its operation in two ways they are testing and training.

STEP-5 The tested and trained data extracts its features using feature extracted method.

STEP-6 Both SVM and ELM will classify the data based on extraction.

STEP-7 At last data is evaluated and classified.

Support Vector Machine (SVM)

The supervise machine learning technique in machine learning algorithms is support vector machine. Classifier is obtained when labeled data is trained in support vector machine [8]. After classification, the labeled data is classified into different number of classes. So, this is nothing but one dimensional (1D) classifier in support vector machine (Fig. 2).

Two-dimensional classifier is nothing but a classifier within 2D space.

Three-dimensional classifier is nothing but a classifier within 3D space.

Four-dimensional classifier is nothing but a classifier within 4D space.

Table 1 shows the comparison of accuracy, precision and $F1$ score of web-based intelligent diagnosis system for SVM, Naïve Bayes, and decision tree. Hence, compared with naïve Bayes and decision tree, SVM of web-based intelligent diagnosis system gives effective outcome.

Figure 3 shows the accuracy of web-based intelligent diagnosis system.

Compared with Naïve Bayes and decision tree, SVM of web-based intelligent diagnosis system improves the accuracy in very effective way.

Figure 4 shows the precision and $F1$ score of web based intelligent diagnosis system.

Compared with naïve Bayes and decision tree, SVM of web-based intelligent diagnosis system improves the accuracy in very effective way.

Table 1 Accuracy of web based intelligent diagnosis system

S. No.	Technique	Accuracy (%)	Precision (%)	$F1$-score (%)
1	SVM	94	97	91
2	Naïve bayes	68	54	73
3	Decision tree	43	395	58

Fig. 3 Comparison of accuracy
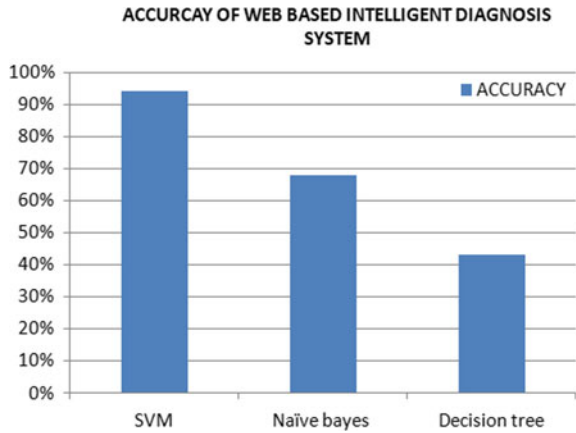
Fig. 4 Comparison of precision and recall
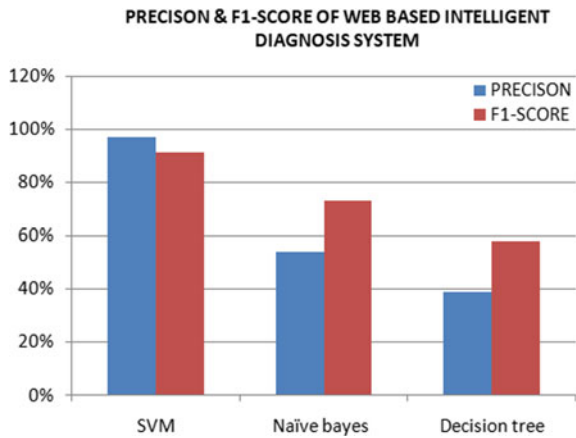
4 Conclusion

Hence in this paper, machine learning techniques to web-based intelligent learning diagnosis system was implemented. The entire flowchart performs its operation in three phases. In first phase, the dataset is investigated based on module and captures the data from dataset. In next phase, the quality of data is enhanced and at last in third phase the noise is deleted in the data. From results, it can be observed that accuracy, reliability, precision, recall, $F1$ score, and Mean gives effective outcome compared with Naïve Bayesian classifier.

References

1. Masood S, Rai A, Aggarwal A, Doja MN, Ahmad M (2018) Detecting distraction of drivers using convolutional neural network. Pattern Recogn Lett
2. Masood S, Srivastava A, Thuwal HC, Ahmad M (2018) Real-time sign language gesture (word) recognition from video sequences using CNN and RNN. In: Intelligent engineering informatics. Springer, Singapore, pp 623–632
3. Masood S, Thuwal HC, Srivastava A (2018) American sign language character recognition using convolution neural network. In: Smart computing and informatics. Springer, Singapore, pp 403–412
4. Thabtah F (2018) Machine learning in autistic spectrum disorder behavioral research: a review and ways forward. Inf Health Soc Care 1–20
5. Thabtah F, Kamalov F, Rajab K (2018) A new computational intelligence approach to detect autistic features for autism screening. Int J Med Informatics 117:112–124
6. Vaishali R, Sasikala R (2018) A machine learning based approach to classify Autism with optimum behaviour sets. Int J Eng Technol 7(4):18
7. Li B, Sharma A, Meng J, Purushwalkam S, Gowen E (2017) Applying machine learning to identify autistic adults using imitation: an exploratory study. PLoS ONE 12(8):e0182652
8. Thabtah FF (2017) Autistic spectrum disorder screening data for adult. https://archive.ics.uci.edu/ml/machine-learningdatabases/00426/
9. Thabtah F (2017) Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment. In: Proceedings of the 1st international conference on medical and health informatics. ACM, pp 1–6
10. Kharazmi E, Försti A, Sundquist K, Hemminki K (2016) Survival in familial and non-familial breast cancer by age and stage at diagnosis. Eur J Cancer 52:10–18
11. McCarthy AM, Yang J, Armstrong K (2015) Increasing disparities in breast cancer mortality from 1979 to 2010 for US black women aged 20 to 49 years. Am J Public Health 105(S3):S446–S448
12. Bone D, Lee C-C, Black MP, Williams ME, Lee S, Levitt P, Narayanan S (2014) The psychologist as an interlocutor in autism spectrum disorder assessment: insights from a study of spontaneous prosody. J Speech Lang Hear Res 57(4):116

Phishing Website Detection Based on Hybrid Resampling KMeansSMOTENCR and Cost-Sensitive Classification

Jaya Srivastava and Aditi Sharan

Abstract In many real-world scenarios such as fraud detection, phishing website classification, etc., the training datasets normally have skewed class distribution with majority (e.g., legitimate websites) class samples overwhelming the minority (e.g., phishing websites) class samples. The machine learning algorithms assume balanced class distributions and are biased towards the majority (uninteresting) class ignoring the minority (interesting) class (es). For handling class imbalance, researchers have proposed solutions both at the (i) data-level and (ii) algorithm-level. In this study we propose a dual approach for handling class imbalance in phishing website classification both at the data and algorithm. We propose a novel hybrid resampling approach KMeansSMOTENCR which balances the dataset by first oversampling the minority class using KMeans Synthetic Minority Oversampling Technique (KMeansSMOTE) (Douzas et al. in Inf Sci 465:1–20, 2018 [1]) followed by Neighborhood Clearing Rule (NCR) (Laurikkala in AIME, LNAI 2001. Springer, Berlin, pp 63–66, 2001 [2]) under sampling technique as the data cleaning approach to take care of the possibility of synthetic minority class samples invading the majority class samples. Finally, we employed Cost-Sensitive Random Forest (CS-RF), Cost-Sensitive Extreme Gradient Boosting (CS-XGB), Cost-Sensitive Support Vector Machine (CS-SVM), and Cost-Sensitive Logistic Regression (CS-LR) classifiers as algorithm-level balancing approach. We evaluated the performance of CS-RF, CS-XGBoost, CS-SVM, and CS-LR classifiers on (i) Original-(Imbalanced), (ii) NCR-(Balanced), (iii) KMeansSMOTE-(Balanced), and (iv) KMeansSMOTENCR-(Balanced) datasets. In Sect. 4 Result and Discussion we demonstrate that the highest ROC_AUC, F1 and GMean are obtained from our proposed method which outperforms the other three. To the best of our knowledge and belief our novel hybrid resampling approach 'KMeansSMOTENCR' has not been published in the existing studies as of now.

J. Srivastava (✉)
Indian Institute of Technology Delhi, New Delhi, India
e-mail: jaya@iitd.ac.in

A. Sharan
Jawaharlal Nehru University (JNU), New Delhi, India
e-mail: aditisharan@mail.jnu.ac.in

Keywords KMeansSMOTE · NCR · KMeansSMOTENCR · Cost-sensitive learning · Random forest · Extreme gradient boosting · XGBoost · Support vector machine · Logistic regression · RF · XGB · SVM · LR · Class imbalance · Data-balancing · Algorithmic-balancing

1 Introduction

In real-world scenarios where anomaly detection is crucial such as fraud detection, electricity pilferage, rare disease diagnosis, phishing website detection, etc., the training datasets suffer from severe class imbalance. But, the conventional machine learning algorithms assume balanced class distributions and does not take into account the prevalent class imbalance in the training datasets resulting into biased and inaccurate predictive models. To address this challenging issue several data balancing [3, 4] as well as algorithm balancing [5, 6] techniques have been proposed in the literature. However, each of such approaches have their advantages and disadvantages and research in handling class imbalance remains an active area of research.

In this study we propose a novel hybrid resampling algorithm KMeansSMO-TENCR based on the existing KMeansSMOTE [1] and NCR [2] methods. Additionally, Cost-sensitive classifiers, namely, CS-RF, CS-XGB, CS-SVM, and CS-LR have been employed and their performance evaluated using ROC_AUC, F1 and GMean. We list below the three important contributions of this study. First, our proposed novel hybrid resampling KMeansSMOTENCR method can be used as a data preprocessing step in handling class imbalance at the data level not only in the domain of phishing website classification but in other domains as well. Second, the proposed method can be used to improve the predictive capability of the cost-sensitive classifiers such as CS-RF, CS-XGB, CS-SVM, and CS-LR that are employed in this study. Thirdly, the best performing model using KMeansMOTENCR is CS-XGB model (ROC_AUC: 99.15%), followed by CS-RF (ROC_AUC: 98.88%), then CS-LR (ROC_AUC: 93.52%) and CS-SVM (ROC_AUC: 93.18%) the least.

The rest of this paper is structured as follows. In Sect. 2 Methodology, we describe the dataset, class imbalance handling at both the data and algorithm level, our proposed 'KMeansSMOTENCR Cost Sensitive Classification' model and the performance metrics used for evaluation purposes in this study. In Sect. 3 The Experimental Setup we mention about the platform used by us to conduct our experiments. Whereas the experimental analysis of our results is presented in Sect. 4 Results and Discussion. Finally, we conclude our findings in Sect. 5 of Conclusion.

Table 1 UCI ML phishing website dataset class distribution

Total samples	Total features	No. of phishing websites	No. of legitimate websites
11,055	30	4898	6137

2 Methodology

2.1 Dataset Description

In this study, we have used publically available benchmark Phishing Website dataset from UCI Machine Learning Repository [7]. The dataset consists of 11,055 samples (rows) and 31 features (columns). Table 1 summarizes the Total Samples, Total Features, No. of Phishing Websites, and No. of Legitimate Websites.

2.2 Class Imbalance Handling

In this study, we have used dual approach for handling class imbalance (A) Data Balancing and (B) Cost-sensitive Learning as discussed below:

Data Balancing. In this study, we explored two existing data resampling methods (i) Neighborhood Clearing Rule (NCR) under sampling [2] and (ii) KMeansSMOTE oversampling [1]. Based on our findings, we propose our novel hybrid resampling method the KMeansSMOTENCR which is a combination of KMeansSMOTE and NCR. Using these three data-balancing techniques, i.e., (i) NCR (ii) KMeansSMOTE, and (iii) KMeansSMOTENCR we generated three class balanced datasets from the Original (Imbalanced) dataset for comparative evaluation purposes.

Cost-Sensitive Classification. Most classification algorithms inherently assume that miss-classifications costs are same [1]. This assumption falls flat when the dataset is imbalanced because misclassifying a minority (interesting class), e.g., a Phishing Website as Legitimate one has far serious consequences than miss-classifying a majority class, i.e., a Legitimate Website as Phishing Website because a Legitimate Website would remain harmless.

Cost-sensitive classification takes misclassification cost into account when training a model [5]. In this study, we have used four Cost-Sensitive classifiers, namely, Cost-Sensitive Random Forest (CS-RF), Cost-Sensitive XGBoost (CS-XGB), Cost-Sensitive Support Vector Machine (CS-SVM), and Cost-Sensitive Logistic Regression (CS-LR) classifiers. Table 2 summarizes the class weights used by the cost-sensitive classifiers. By setting up of the hyper parameter class_weight = 'balanced', the traditional cost-insensitive classifiers transform into cost-sensitive classifiers. In the "balanced" mode the classification algorithm automatically adjusts and sets class weights to be inversely proportional to the class frequencies in the

Table 2 Class weights computations of the two classes: phishing and legitimate

Class weight (cw$_1$) for phishing websites	Class weight (cw$_2$) for legitimate websites
= 11,055/(2 * 4898) = 1.28522	= 11,055/(2 * 6137) = 0.900684

target label. The formula used for computation of class weight for class C_i is given below:

$$\text{Class Weight of } C_i \ (\text{cw}_i) = \text{Total samples} \div (\text{No. of Classes} * \text{Total samples of } C_i) \tag{1}$$

The No. of classes = 2 in Eq. (1) for the binary classification problems.

2.3 Proposed Model

Figure 1 depicts our novel proposed model 'KMeansSMOTENCR Cost-Sensitive Classification' method.

As is evident from Fig. 1, in our proposed KMeansSMOTENCR hybrid resampling method, we first perform KMeansSMOTE Oversampling [1] on the Original imbalanced dataset to generate new synthetic minority samples. The KMeansSMOTE algorithm first performs KMeansClustering followed by SMOTE to generate the new synthetic minority samples. However, the possibility of KMeansSMOTE introducing noise in terms of the new synthetic minority class samples invading into the existing majority class samples is taken care by noise removal done using neighborhood clearing rule (NCR) [2] under sampling technique. NCR combines both condensed nearest neighbor (CNN) and edited nearest neighborhood (ENN) method as underlying data cleaning techniques. CNN is employed first to remove the redundant or duplicate majority samples. ENN is used to remove to remove noisy or ambiguous data. NCR focuses on quality i.e. unambiguity of the data. After data-balancing we
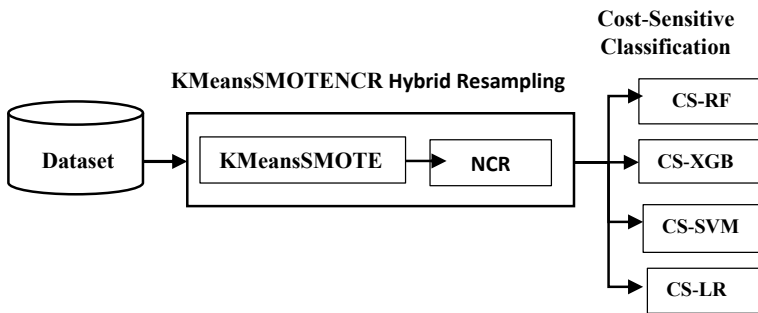
Fig. 1 Proposed 'KMeansSMOTENCR' cost-sensitive classification model

used CS-RF classifier, CS-XGB classifier, CS-SVM, and CS-LR classifier as an algorithmic-balancing technique in place of their traditional counterparts to generate four models.

2.4 Performance Metrics

The performance evaluation metrics used in this study are Area under Receiver Operating Characteristics (ROC_AUC) Curve, F1, GMean and Accuracy. Empirical results show that our novel hybrid resampling technique KMeansSMOTENCR outperforms NCR and KMeansSMOTE approaches.

3 Experimental Setup

All our experiments have been performed on Intel® Core™ i7-4770S CPU @ 3.1 GHZ with 8.00 GB RAM and 64-bit Windows Operation System. The Anaconda Navigator (64 bit) version 4.10.3 with jupyter Notebook version 6.0.3 and Python version 3.6.12 formed the experimental test bed.

4 Results and Discussion

Table 3 presents the comparative evaluation of our four models namely, CS-RF, CS-XGB, CS-SVM ans CS-LR. The highest values of the performance metrics in the last column, namely, KMeansSMOTENCR, our novel proposed method, are highlighted in bold.

Our proposed method KMeansSMOTENCR significantly enhanced the performance of the (i) CS-RF model on the ORIGINAL (Imbalanced) dataset from 97.72% (ROC_AUC) to 98.88% (ROC_AUC), from 97.53% (F1) to 98.86% (F1), from 97.71% GMean) to 98.88% (GMean) and from 97.83%(Accuracy) to 98.91% (Accuracy). Similarly, performance of the (ii) CS-RF model on KMeansSMOTE got enhanced from 97.93% (ROC_AUC) to 98.88% (ROC_AUC), from 97.91% (F1) to 98.86% (F1), from 97.93% (GMean) to 98.88% (GMean) and from 97.93% (Accuracy) to 98.91% (Accuracy) using our proposed method KMeansSMOTENCR. When compared with performance of the (iii) CS-RF model on NCR we got the performance boost from 98.22% (ROC_AUC) to 98.88% (ROC_AUC), from 98.11% (F1) to 98.86% (F1), from 98.22% (GMean) to 98.88% (GMean) and from 98.23% (Accuracy) to 98.91% (Accuracy) using our novel proposed KMeansSMOTENCR method.

Similar results can be inferred for the other three models, i.e., CS-XGB classifier, CS-SVM and CS-LR classifier from Table 3. The performance metrics of these four

Table 3 Comparative evaluation of cost-sensitive classifiers

Cost-sensitive classifier	Performance metric	ORIGINAL dataset (imbalanced) (%)	Data resampling technique		
			KMeans-SMOTE (%)	NCR (%)	KMeans-SMOTENCR (%)
1. CS-RF	ROC_AUC	97.72	97.93	98.22	98.88
	F1	97.53	97.91	98.11	98.86
	GMean	97.71	97.93	98.22	98.88
	Accuracy	97.83	97.93	98.23	98.91
2. CS-XGB	ROC_AUC	96.62	97.32	98.55	99.15
	F1	96.30	97.3	98.47	99.13
	GMean	96.61	97.32	98.55	99.15
	Accuracy	96.74	97.32	98.57	99.16
3. CS-SVM	ROC_AUC	90.50	91.43	92.70	93.41
	F1	89.43	91.29	92.22	93.18
	GMean	90.45	91.42	92.69	93.41
	Accuracy	90.82	91.43	92.78	93.41
4. CS-LR	ROC_AUC	90.90	92.25	92.28	93.52
	F1	89.86	92.22	91.76	93.3
	GMean	90.89	92.25	92.27	93.52
	Accuracy	91.00	92.25	92.36	93.52

models are further plotted in the multiple-bar chart as shown in Fig. 2 (CS-RF), Fig. 3 (CS-XGB), Fig. 4 (CS-SVM), and Fig. 5 (CS-LR). As is evident from Table 3 and these four figures it is clear that our novel proposed 'KMeansSMOTENCR Cost-Sensitive Classification' approach consistently outperformed the other three approaches, i.e., the ORIGINAL, the NCR and the KMeansSMOTE in all the four performance metrics, i.e., ROC_AUC, F1, GMean, and Accuracy in all the four models, i.e., CS-RF, CS-XGB, CS-SVM, and CS-LR. Further, CS-XGB model outperformed the rest, followed by CS-RF, then CS-LR and CS-SVM the last.

5　Conclusion

Handling the inescapable class imbalance prevalent in the real-world datasets plays a significant role in developing unbiased and better predictive models. The major contributions of this study can be summarized into three. Firstly, we have demonstrated that our novel proposed 'KMeansSMOTENCR' method significantly enhanced the Accuracy, GMean, F1 and ROC_AUC of the four cost-sensitive classifier models i.e. CS-RF, CS-XGB, CS-SVM and CS-LR as compared to the (i) ORIGINAL (Imbalanced), (ii) KMeansSMOTE, and (iii) NCR methods. Secondly, we

Fig. 2 Performance of CS-RF classifier
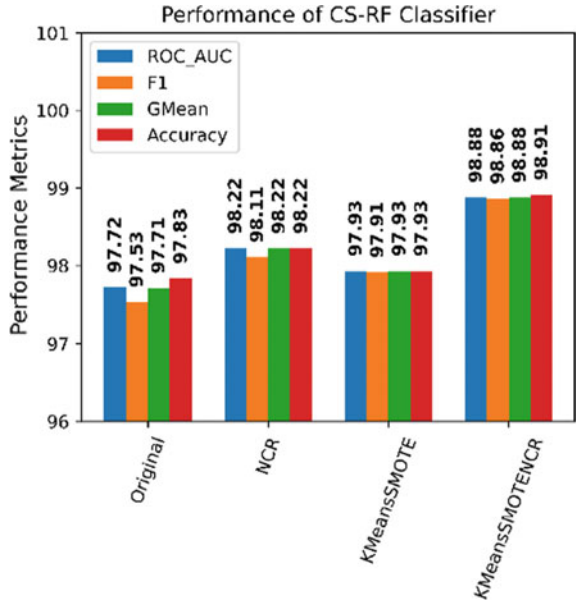
Fig. 3 Performance of CS-XGB classifier
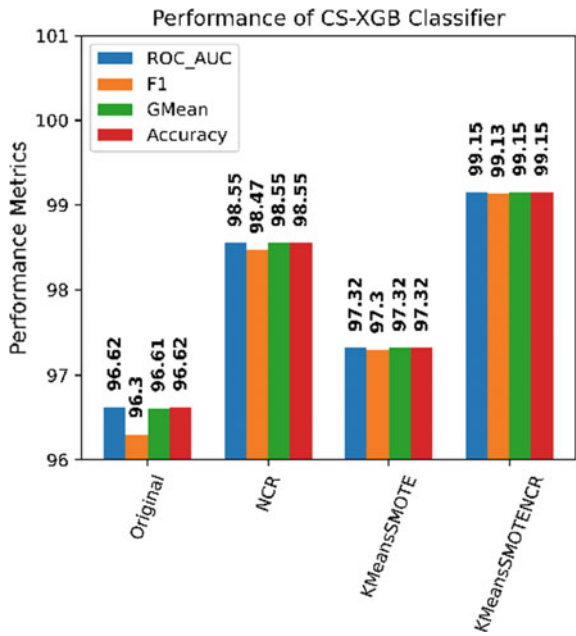
Fig. 4 Performance of CS-SVM classifier
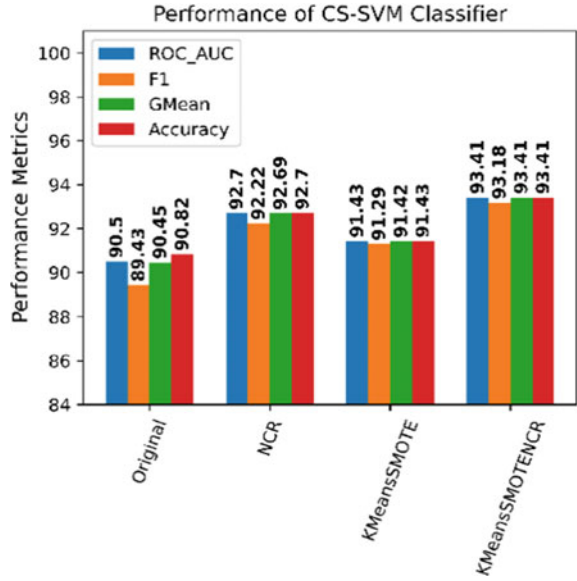
Fig. 5 Performance of CS-LR classifier
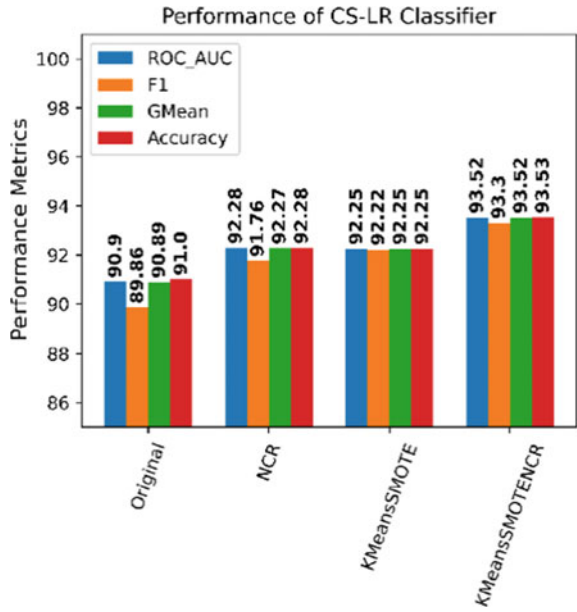
have also demonstrated that a dual combination of data-balancing and algorithmic-balancing solutions can significantly enhance the classification performance. Thirdly, as is evident from the results it can be said that our novel proposed 'KMeansSMO-TENCR Cost-Sensitive Classification' model can easily be used not only in the detection of Phishing Website but in other problem domains as well. In the future, we would like to further extend our research in the direction of proposing more novel and better techniques for handling class imbalance leading to better 'Phishing Website Detection' models which may also be employed in other problem domains as well.

References

1. Douzas G, Bacao F, Last F (2018) Improving imbalanced learning through a heuristic oversampling method based on *k*-means and SMOTE. Inf Sci 465:1–20. https://doi.org/10.1016/j.ins.2018.06.056
2. Laurikkala J (2001) Improving identification of difficult small classes by balancing class distribution. In: AIME, LNAI 2001. Springer, Berlin, pp 63–66. https://doi.org/10.1007/3-540-48229-6_9
3. Pristyanto Y, Dahlan A (2019) Hybrid resampling for imbalanced class handling on web phishing classification dataset. In: 4th international conference on information technology, information systems and electrical engineering (ICITISE). https://doi.org/10.1109/ICITISEE48480.2019.9003803
4. Azari A, Namayanja JM, Kaur N, Misal V, Shukla S (2020) Imbalanced learning in massive phishing datasets. IEEE 978-1-7281-6873-9/20. https://doi.org/10.1109/BigDataSecurity-HPSC-IDS49724.2020.00032
5. Ren Z, Zhu Y, Kang W, Fu H, Niu O, Gao D, Ke Y, Hong J (2022) Adaptive cost-sensitive learning: improving the convergence of intelligent diagnosis models under imbalanced data. Knowl Based Syst. https://doi.org/10.1016/j.knosys.2022.108296
6. Thai-Nghe N, Gantner Z, Schmidt-Thieme L (2010) Cost-sensitive learning methods for imbalanced data. In: International joint conference on neural networks (IJCNN). IEEE. https://doi.org/10.1109/IJCNN.2010.5596486
7. Mohammad RM, Thabtah F, McCluskey L (2014) Predicting phishing websites based on self-structuring neural network. Neural Comput Appl 25(2):443–458. https://doi.org/10.1007/s00521-013-1490-z

A Brain-Inspired Cognitive Control Framework for Artificial Intelligence Dynamic System

Mrutyunjaya S. Yalawar, K. Vijaya Babu, Bairy Mahender, and Hareran Singh

Abstract From agriculture to medical field, each industry is going through the revolution while embracing the artificial intelligence (AI). However, still AI is in its infancy compared to biological brain cognitive abilities. To bridge this gap, a brain-inspired cognitive framework for artificial intelligence dynamic system is presented in this research for synthesizing an artificial brain with cognitive abilities. This framework will share the few of the image key characteristics from multiple kernel image sizes between neurons for converting image data into neuron activation. Then cognitive controller with learning and planning using Bayesian estimator is employed for making it capable of learning and memorizing images. In cognitive controller, an executive learning algorithm can be computed through entropic state processing, followed by predictive planning for setting the stage to policy for acting over the environment thereby global perception action cycle establishment. In perception, improved sparse coding will be utilized with perceptual attention influence to extract relevant data from observables and ignore the irrelevant data. Then to estimate the state, Bayesian algorithm is utilized. Experimental results illustrate that this framework will have potential for incrementally improving itself over the generations based on the performance of system and utilized genetic algorithm.

Keywords Artificial intelligence (AI) · Cognitive controller · Brain-inspired framework · Bayesian algorithm

M. S. Yalawar (✉) · K. Vijaya Babu · B. Mahender
CMR Engineering College, Hyderabad, India
e-mail: muttusy@gmail.com

K. Vijaya Babu
e-mail: k.vijaybabu@cmrec.ac.in

H. Singh
GNIOT, Greater Noida, India

1 Introduction

Artificial intelligence is playing an enormous role in the operational and scientific research area. The ability of an artificial intelligence program is to collect data and information to solve complex problems being on an outstanding scale [1]. In the future, these intelligent program incorporated in machines may or could replace human instincts in many advanced areas. Artificial intelligence is the study and incorporation of intelligent processing inside a system or a machine [2]. These artificial intelligent machines could learn things on their own, adapting its surroundings. It could basically gather and process information more efficiently and effectively than a human brain, terming this program an "electronic brain". The decision that the e-brain in favor is based on the contrary that, it first takes an option from the list of the options available [3]. Then it computes and process the decision based on the positive aspects and the negative aspects from more than a million probability of the decision taken could cause a conflict or chaos, and takes the most favored option as its final outcome. Thus, e-brain being the most complex and the advanced version of artificial intelligence, that can be incorporated in any digital device, it could be either a normal smart phone or any machine [4]. Data in various form factors are analyzed and preprocessed to make decision that could sort by itself into the best possible outcome state. Artificial intelligent machine could improve from what it is now with the enormous ability to improve its decision-making capability and natural language processing factor [5]. Artificial intelligent programs could enable in developing and making a new complicated versions of artificial intelligence system that could understand the human ideology of thinking and could even communicate as one among us, hence, terming it as the electronic brain [6]. The brain of human is a most powerful information processing machine and no equal is there for it when comes to two challenging problems: environment perception and action over the environment (World) [7]. The cognitive dynamic system (CDS) idea is inspired from brain. Hence, not surprisingly by embracing the CDS, the classical entities are outperformed in the orders of magnitude on cognitive control and cognitive radar [8]. Thus a novel framework and paradigm is presented, namely brain-inspired cognitive control framework to help in building the intelligent systems inspired from brain, that will grow, scale, evolve and adapt to the world, confined efforts are required for defining the structural characteristics to any targeted cortical area for simulation. Based on the biological brain's physiology and anatomy segregation, this presented model enables the simplified yet scalable design process.

2 Literature Survey

Steunebrink and Schmidhuber [9] said that Gödel Machine can be viewed as a program that contains two parts. First part is called as solver that might be any problem-solving program. The solver is a reinforcement learning (RL) program

interacts with certain external environments for the clarity of presentation. This can provide a convenient way for determining the utility (utilizing the reward function of RL) which is a significant concept later on. But generally no constraints can be placed over the solver. Gödel Machine second part is called as searcher, and it is a program that trying for improving the whole Gödel Machine (includes searcher) in a provably optimal manner.

Achler in [10] presented an interesting, pragmatic AGI intelligence measurement technique inspired from the general techniques, in the sense that explicitly balancing the system effectiveness during problem-solving with its solutions compactness. It is similar like a basic mechanism in evolutionary program learning in which one will utilize the fitness function having an "Occam's Razor" compactness term and accuracy term.

Laird et al. [11] presented a requirements list to the intelligence level of human from designers' standpoint of cognitive structures. Mostly, this work involved the architecture of cognitive SOAR that is pursued from the AGI perspective and from accurately simulating human cognition perspective.

Bach et al. [12] have elegantly characterized it in quest for creating the "synthetic intelligence". It can also determine that the research communities are working toward the goals of AGI in the labels "cognitive architecture", "BICA" (biologically inspired cognitive architecture), "natural intelligence", "computational intelligence" and several others. These labels are launched with few underlying purposes and contain a particular concepts collection and associated techniques; each one relates to some perspectives or perspectives family.

Kurzweil et al. [13] have utilized "narrow AI" term for indicating the systems creation which performs specific "intelligent" behaviors in particular contexts. To a narrow AI system, when anyone changed the behavior specification or context to a little bit, certain levels of human reconfiguration or reprogramming are essential for enabling the system to retain intelligent levels.

Modha's IBM "Cognitive Computation Project", aimed for "reverse engineering the structure, function, dynamics and behavior of the human brain, and then delivering it in a small compact form factor consuming very low power that rivals the power consumption of the human brain". Modha's team best published accomplishment is neural network simulation (at a particular accuracy level), the "cortex of a cat" size with 109 neurons and 1013 synapses (Frye [14]).

EPIC (Rosbe et al. [15]), a cognitive architecture is presented that is aimed to capture the perceptual of human, motor and cognitive activities via many interconnected processors which would be working in parallel. This system can be controlled by the rules of prediction to cognitive processor, a set of perceptual (tactile, auditory and visual) and motor processors operated on symbolically coded features instead of raw sensory information. It is attached to SOAR for learning, planning and issue solving.

3 A Brain-Inspired Cognitive Control Framework for AI

This brain-inspired cognitive control AI system framework is designed as shown in Fig. 1 for receiving the raw image data from different input portals and having that data is readily available for feeding to cortical pathways in neuron simulations form. For understanding the output of artificial brain either by computer or by human, the activities of neuronal should be processed and passed to relevant output devices. The biological brain collects' the data from peripheral nervous system which is connected to several receptors like mechanoreceptors, chemoreceptors, thermoreceptors and photoreceptors. In the same way, here the brain-inspired AI system, every input processing unit/output processing unit (IPU/OPU) can differ that how data can be extracted, encoded and presented for first cortical layer as a specific functional critical pathway part.

In this framework, a designated AI system is there for visual data that will convert the given object image in to its primary elements like brightness and color by maintaining the arrangements of pixel, same as to the structure of retinal information processing. The human eye will collect the visual data by photoreceptors and forward it via cell assortment before passing it to thalamus through ganglion cells. In this framework, the brain-inspired cognitive control AI system plays the similar role like retina and remaining processing activities can be handed off in neuronal simulation form to neural processing unit represents the cortical pathways.

In brain-inspired cognitive control AI system design, two primary principles are applied that can be inspired from nature. First principle depends on how the projection of neuron maintained in their fidelity position against their neighborhoods since they projected from retina to LGN and for striating cortex. Second one based on how horizontal cells with the dendrites spanning side paths collecting the data from

Fig. 1 Brain-inspired cognitive control AI system framework
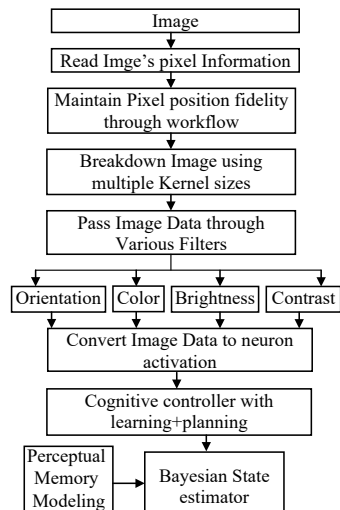
neighborhood neurons. These principles can be applied by leveraging the most popular method in image processing which includes a kernel convolution matrix through where smaller size matrix is utilized for scanning the image when applied to the filters. This technique helps for dividing the images into primary elements like contrast, brightness, color and orientation. These primary parameters sensitivity causes neuron activation at visual cortical pathway beginning.

The memory decoding process of back to OPU needs appropriate neuron wiring from region of memory to OPU that will be driven through genome. Secondly, the original neural signal decoding achieved from its relevant memory unit to real physical signal required to be interacting with the system which is outside the artificial brain. To prove this concept, only cognitive controller OPU is implemented. The CDS system in its common form contains two most dominant physical units: controller and preceptor. This system works in a synchronized and self-organized way through building over the primary cognitive neuroscience principles. During the perception action period, the cycle starts with the processing of incoming world observables follow by feedback data of world send to the controller through the preceptor for setting the stage to controller for acting over the environment. Generally, this leads to variations in the observables of environment, and this will set the stage to second action cycle of preceptor and it goes on. The CDS distinctive cyclic behavior will be continued till it reach a point at which information of environment will be very small in practical value and that environment can be assumed as stationary.

Basically, the perceptual memory function is solving the issue of source separation that is attained by the extraction of corresponding information from the observables of environment which retains relevant data and ignores irrelevant and unwanted data. This issue is solved by continually learning from observables in cyclic basis. The function of preceptor of to model the observables behavior followed by estimation of Bayesian State. In addition, perceptual actions depend on local perception cycle to their corresponding impacts over cognitive functions performed in executive and perceptual memories with enhanced performance at end results.

In neuroscience, rooted sparse coding plays a vital role while addressing the perception function of cognitive. The sparse coding problem solution relies on two components. For solving the source separation issue (i.e., separation of relevant data from irrelevant) in CDS preceptor, strong arguments can be made. The sparse coding algorithm is well posed, sufficient data can be there in observables, subjected for the provision that the signal to noise ratio cannot be very less. Bayesian data is recognized as inherent characteristics of cognitive perception and Bayesian filtering specific case is information filtering. Hence, the sparse coding performance can be improved under the perceptual attention influence by the information filtering utilization in Bayesian. This Bayesian algorithm plays a vital role in decision-making to the action which is taken by the controller over environment. The preceptor entropic state model can be defined formally through the equation of entropic state:

$$H_k = \phi(p(x_k|z_k)) \tag{1}$$

where H_k is entropic state during cycle k in accordance with the state posterior $p(x_k|z_k)$ in Bayesian sense that can be computed in preceptor. As that H_k is the preceptor state and ϕ is a quantitative measure, which can be a Shannon's entropy.

In the sense, the cognitive control is seen as CDS overreaching function on the function that it can perform in the system. In particular, the cognitive control function may define as for controlling the entropic state (i.e., preceptor state), and the estimated state of target will be expected to be reliable and continue across the time. For elaborating more, the cognitive control includes two process, namely predictive planning and executive learning, and aim of these processes is policy formation. The policy would be a cognitive actions probability distribution carried out on environment at cycle $k + 1$, conditioned in the cognitive activity chosen at present cycle k.

Executive Learning in Cognitive Control
The function of value-to-go can be defined, denoted by $J(c)$ to the cognitive controller as

$$J(c) = \mathbb{E}^\pi[r_{k+1} + \gamma r_{k+2} + \gamma^2 r_{k+3} + \ldots |c_k = c] \tag{2}$$

where $\gamma \in [0, 1)$ indicates a discount water which is continually reduces the future action effect and $\mathbb{E}^\pi$ refers the expected value of operator to that expecting value is computed by the policy distribution π_k. It is significant to notice that the $J(c)$ function denoted in (2) is stateless.

Let $\mathscr{R}(c) = \mathbb{E}^\pi[r_{k+1}|c_k = c]$ defines the expected immediate rewards at $k + 1$ cycle to present chosen actions at cycle k. The n it will prove that the function of value-to-go referred in (2) meets the recursions:

$$J(c) = \mathscr{R}(c) + \gamma \Sigma_{c'} \pi_k(c, c') J(c') \tag{3}$$

For having a recursive algorithm, Eq. (3) recursion can be expressed in below update format:

$$J(c) \leftarrow \mathscr{R}(c) + \gamma \Sigma_{c'} \pi_k(c, c') J(c') \tag{4}$$

While the learning algorithm is considered because it is typical during neural computation a learning parameter is introduced and it is referred as $\alpha > 0$ on the basis of which might be expressed as:

$$J(c) \leftarrow J(c) + \alpha[\mathscr{R}(c) + \gamma \Sigma_{c'} \pi_k(c, c') J(c') - J(c)] \tag{5}$$

Now, the function of value-to-go will be updated form one cycle to next. The update in (5) is known as executive learning algorithm.

For summarizing, this algorithm in cognitive control is derived with the exploitation of two common thoughts rooted in CDS:

(1) The cyclic flow of directed data passes in cyclic way the first principle, i.e., the action cycle of perception.
(2) The entropic state that resides in preceptor in two state scheme.

Predictive Planning in Cognitive Control

The execution learning algorithm depends on a selected action influence on preceptor through the environment in every global action cycle of perception. This influence itself manifesting in global cycle entropic state. Along with the process of executive learning, cognitive controller might also benefited by using future entropic state successive predictions. Hence, it is required to be consider another learning process which is similar to executive learning algorithm; however, this time to predict entropic states pertaining for certain hypothesized activities: In the second learning process, the value-to-go function depends on entropic reward predicted values is known as planning. It can also define as planning constituted the second intrinsic learning process in the cognitive control.

By taking the predictive planning in to account, it may following a feed forward link which is desirable in CDS design. An internally composite cycle is resulted by adding the feed forward link, that will bypass the environment and future rewards prediction can be permitted. The information of feed forward in every internal composite cycle so a hypothesized future action that can be chosen to plan a stage. With this hypothesized action, relevant entropic state will be computed in preceptor thereby the process of planning will be initiated. For cognitive control, complete algorithm is involving the combine usage of predictive planning and executive learning.

4 Results

For demonstrating this framework success and feasibility, visual system simple simulation is implemented that is confined for single vision in gray scale and is only capable to static images processing. The implementation of proof-of-concept is evolving via 20 generators. An interested area is to proving the artificial brain designed by brain-inspired cognitive AI framework which has improving capabilities itself over generations. For evaluating these improvements, fitness function is referred as

$$\text{Fitness} = \frac{1}{1 + e^{(-t+a)}} \times \frac{c}{t} \tag{6}$$

Here, c indicates the number of times that the artificial brain is capable for a stimulus identification, a is defines threshold of activity and t is total number of times that the stimulus is exposed to brain.

Fig. 2 Artificial brain's fitness improvement over generations
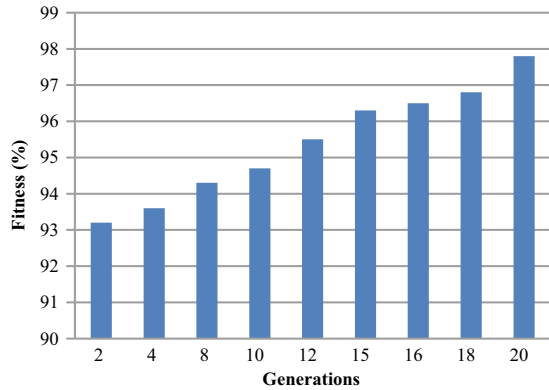
A baseline genome is developed that demonstrates the different cortical regions limited physiological characteristics, growth rules, anatomical properties and brain is set for evolving over generations with two sets of MNST digits from 0 to 9. Every generation has been gone via self-assessment and self-learning processes as a consequence value of fitness is calculated and associated to the genome which is given. From Fig. 2, it is clear that the brain fitness has significantly improved in last generations compared to earlier generations and totally positive trend.

Here, the described experiments are intended for demonstrating how CDS will control the directed flow of information satisfactorily in severe disturbance presence. This experiment is being concerned with falling object tracking in non-stationary environmental space at which falling object will experience a disturbance. The disturbance might be appeared by strengthening the system noise power in state space model and again reduced to its beginning setting; all of these are completed in prescribed time period in a way which is represented in Fig. 3.

To the first experiment on risk control, only a process is followed, which is built upon multilayer preceptors rooted in the computations of neural. For elaboration,

Fig. 3 Overall system noise for duration from 75 to 150 cycles
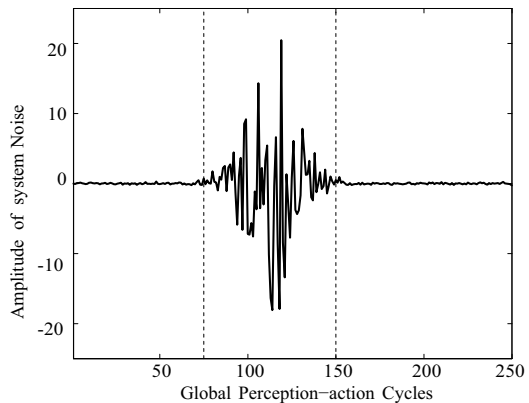
Fig. 4 Learning curve analysis of cognitive control system
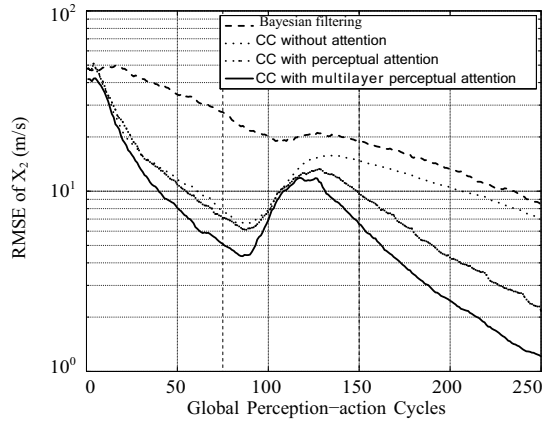
adopted the probability distribution of system noise on a cyclic basis (cycle-to-cycle) through observing the past residual errors window which is available. This can be simplified by assuming the residual errors amplitude spectrum as the filter can be expected to be flat to a longer error sequence. Meanwhile, if the filter is set up with the parameters of imprecise approach, then residual errors witness won't be hold. Hence, in this experiment, an attempt is made for adapting the noise distribution of system by successive residual errors amplitude spectra analysis. Due to this, a multilayer perceptron is utilized that will take the successive residual errors amplitude spectra window as its input. Then it will identify the extent where present system noise variance value from the must be decreased or increased during iterations.

While computing the experiment, it should be vital to notice that a disturbance with 30 dBs at peak point cab be applied through increasing the system noise magnitude computed with reference to the segment which is attributed for stationary system noise, apparently this would be a severe disturbance for dealing with it. While considering this disturbance, the experiments results are involved with various tracking structures are drawn in Fig. 4.

The Bayesian filter works as target tracker by its self and can be complete failure as soon as disturbance initiates for dominating the observables. Without attention, the target controller nicely tracks the target. However, as the disturbance is introduced, the trajectory of target moves upward until it reaches the disturbance tip. After that, it would loss control and trailed in a way which is similar as CKF during first case. Then with perceptual attention, the cognitive controller follows trajectory condition, however, the perceptual attention supports the controller for starting the recovery from disturbance after reaching the tip of disturbance efficiently. At last, not surprisingly, completely equipped cognitive controller performs better among rest of four trackers because of the combined usage of executive and perceptual attention.

5 Conclusion

A brain-inspired cognitive control framework for artificial intelligence dynamic system is presented in this paper which has effective visualization capabilities and cognitive abilities than other brain simulation frameworks. In this framework, differentiating factors are scalability and evolvability enabled by selected architecture basing on new sparse encoding method inspired from the process of neuroembryogenesis in the embryo of human. It is recognized that entropic state and perception play a vital role itself in designing the framework for cognitive control. Meanwhile, the cognitive control in CDS is to control the preceptor state that defines the state of entropic. From the result analysis, it is demonstrated that an artificial visual cortex created from a genome simulates the human brain ventral visual pathway which is capable for learning, recalling and memorizing the digits in real time read from the database MNIST with higher fitness values and with latest generations.

References

1. García DH, Adams S, Rast A, Fang TH, Zeng Y, Zhao F (2021) Brain inspired sequences production by spiking neural networks with reward-modulated STDP. Front Comput Neurosci 15:8
2. Rodriguez J (2019) Beyond neurons: five cognitive functions of the human brain that we are trying to recreate with artificial intelligence
3. Furber WS, Cangelosi A (2018) Visual attention and object naming in humanoid robots using a bio-inspired spiking neural network. Robot Auton Syst 104:56–71
4. Fitzgerald JE et al (2017) Artificial nose technology: status and prospects in diagnostics. Trends Biotechnol 35(1):33–42
5. Alex F, Andrew Z, Bullmore ET (2016) Fundamentals of brain network analysis. Academic Press, pp 1–35
6. Goertzel B (2014) Artificial general intelligence: concept state of the art and future prospects. J Artif Gen Intell 5(1):1–48
7. Sumari ADW (2013) A new model of information processing based on human brain mechanism: toward a cognitive intelligent system. In: Proceedings of the 1st conference on information technology computer and electrical engineering (CITACEE 2013). Diponegoro University, 16 Nov 2013, pp 56–61
8. Sumari ADW, Ahmad AS, Wuryandari AI Sembiring J (2012) Brain-inspired knowledge growing-system: towards a true cognitive agent. Int J Comput Sci Artif Intell (IJCSAI) 2(1):26–36
9. Steunebrink BR, Schmidhuber J (2012) Towards an actual Gödel machine implementation: a lesson in self-reflective systems. In: Wang P, Goertzel B (eds) Theoretical foundations of artificial general intelligence. Atlantis thinking machines, vol 4. Atlantis Press, Paris
10. Achler T (2012) Towards bridging the gap between pattern recognition and symbolic representation within neural networks. In: Workshop on neural-symbolic learning and reasoning, AAAI-2012
11. Laird JE, Wray R, Marinier R, Langley P (2009) Claims and challenges in evaluating human-level intelligent systems. In: Proceedings of the second conference on artificial general intelligence, pp 91–96
12. Bach J (2009) Principles of synthetic intelligence PSI: an architecture of motivated cognition, vol 4. Oxford University Press

13. Kurzweil R (2005) The singularity is near: when humans transcend biology. Penguin
14. Frye J, Ananthanarayanan R, Modha DS (2007) Towards real-time, mouse-scale cortical simulations. CoSyNe: computational and systems neuroscience. Salt Lake City, Utah
15. Rosbe J, Chong RS, Kieras DE (2001) Modeling with perceptual and memory constraints: an EPIC-soar model of a simplified enroute air traffic control task. Report. SOAR Technology Inc

Meandered Shape CPW Feed-Based SIW Slot Antenna for Ku-Band Applications

Sai Padmini Vemu, S. Mahaboob Basha, and G. Srihari

Abstract In this article, the antenna performance of a novel cavity-backed triangular slot-based substrate integrated waveguide (SIW) antenna is explored by gain, S-parameters, and radiation pattern. SIW approaches are low-cost, small-scale, and easy to integrate into a planer circuit. The planned antenna construction is built on a Roger RT 5880 substrate of 1.57 mm thickness, 2.2 dielectric constant, and 0.0009 as tangent loss. The suggested antenna is fed by a tapered, meandering-shaped CPW-to-SIW transition that has superior electrical performance. CST Microwave Studio was used to design this antenna. Simulated findings reveal that the suggested antenna may achieve gain and directivity of 3.107 dB and 8.316 dBi at 16.461 GHz center frequency, and 5.644 dB gain and 10.14 dBi directivity at 17.636 GHz, respectively. The proposed antenna is small, has a plain structure, and can be employed in a diversity of Ku band applications.

Keywords Substrate integrated waveguide (SIW) · Co-planar waveguide (CPW) · Defected ground structure (DGS) · Ku-band · Satellite application · Cavity backed triangular slot antenna

1 Introduction

Due to their high gain, cavity-backed antennas (CBAs) have made remarkable progress in space communication in recent years, and many academics have researched them intensively. In general, traditional CBAs are designed with large metallic chamber toward eliminate backside radiation [1]. This prevents such

S. P. Vemu
Vignan Institute of Engineering for Women, Vizag, India
e-mail: Padmini10877@view.edu.in

S. Mahaboob Basha · G. Srihari (✉)
Sree Vidyanikethan Engineering College, Tirupati, India
e-mail: srihari.g@vidyanikethan.edu

S. Mahaboob Basha
e-mail: Mahaboobbahsa@vidyanikethan.edu

antennas from being used in contemporary small wireless systems, which favor light mass, low-profile, and planar integration [2]. For construction of planar CBA structure, a relatively new approach identified as "substrate integrated waveguide" (SIW) technology has arisen to meet these necessities for contemporary wireless communication systems.

The supported cavity is built entirely on a PCB substrate, with a metalized via array running length of the substrate and two metallic layers on top and bottom surface. While certain parameters are met, the attenuation constant will be little adequate that leaking from two close vias can be ignored. High radiation presentation of conservative cavity-backed antennas, such as large gain, less back-lobe, and less cross-polarization level, may be maintained with our innovative cavity-backed antennas [3]. SIW-based CBAs have all of the benefits of traditional CBAs, including small transmission loss, high-quality factor, and easy to manufacture by normal printed-circuit-board processes. As a result, mass production becomes more efficient [4].

This work presents a triangular slot DGS antenna for satellite applications with a substrate integrated waveguide supplied through meander CPW. The purpose of this cavity-SIW was to increase gain while keeping superior radiation behavior. Simulation result shows that proposed antenna may give 5 dB gain boost while achieving a reflection coefficient < -10 dB at dissimilar band frequencies (12.67 and 14.56 GHz).

2 Antenna Configuration and Theoretical Analysis

SIW is a structure reported in [5] that is built on a dielectric substrate through episodic linear array of tinny vias [6, 7]. SIW demonstrates highly potential performance enhancement and appealing downsizing methodologies in microwave component design.

Effects of spacing parameter, p, and diameter, d (see Fig. 1) investigated [8] on the freedom from any radiation losses, and the design principles for SIW of width w were set as follows.

$$\frac{p}{d} < 2 \quad \text{and} \quad \frac{d}{w} < 1.5 \tag{1}$$

The center-to-center width and length (W_{SIW} and L_{SIW}) between rows of vias, diameter (d) of vias, and spacing (P) between vias are SIW cavities' design parameters. In conventional metallic cavity filled by similar dielectric and propagation properties of the SIW cavity are remarkably comparable to those of a typical metallic cavity. TE_{120} is the dominant mode stimulated by a rectangular cavity. The SIW cavity's resonant frequency related by its width and length effectively [9]
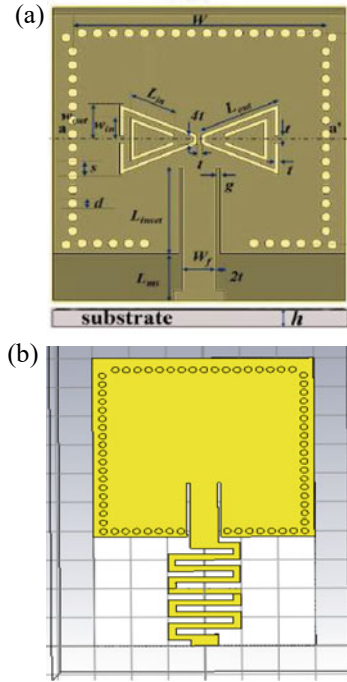
Fig. 1 Proposed antenna design **a** front view with dimension parameters **b** back view geometry of proposed design ($L = 29$ mm, $W = 27.3$ mm, $w_s = 30$ mm, $l_s = 50$ mm, $h = 1.57$ mm, $d = 1$ mm, $s = 1.5$ mm, $t = 0.035$ mm, $w_f = 3.8$ mm, $w_0 = 5$ mm, $l_{ms} = 19$ mm, $g = 0.4$ mm, $l_{inset} = 6$ mm, $l_{out} = 9$, $\theta = 0°$, and $\varphi = 90°$)

$$f_r(\text{TE}_{120}) = \frac{C}{2\sqrt{\varepsilon_r}}\sqrt{\left(\frac{1}{W_{\text{eff}}}\right)^2 + \left(\frac{2}{L_{\text{eff}}}\right)^2} \qquad (2)$$

where $\varepsilon_r =$ permittivity, $L_{\text{eff}} =$ effective length, and $W_{\text{eff}} =$ effective width of SIW, respectively. The effective width and length can be approximated as follows:

$$L_{\text{eff}} \text{ or } W_{\text{eff}} = (L \text{ or } W) - 4.32\frac{r^2}{p} + 0.4\frac{r^2}{(L \text{ or } W)} \qquad (3)$$

To mix SIW and CPW technologies, CPW transitions are generally necessary. Through a tapered CPW section (Fig. 1), CPW conductor is connected to top wall of SIW, and the CPW ground plane is connected to a bottom wall of SIW, resulting in a dense, single layer transition connecting the CPW and SIW. The taper conversion is accountable for ensuring impedance matching between the feeding line and the SIW cavity.

With Roger RT 5880 substrate, 1.57 mm thickness, 2.2 dielectric constant, and 0.0009 tangent loss (dimensions $Ls \times Ws \times hs$), the planned antenna is designed. It

Fig. 2 Proposed antenna design **a** front view **b** side view
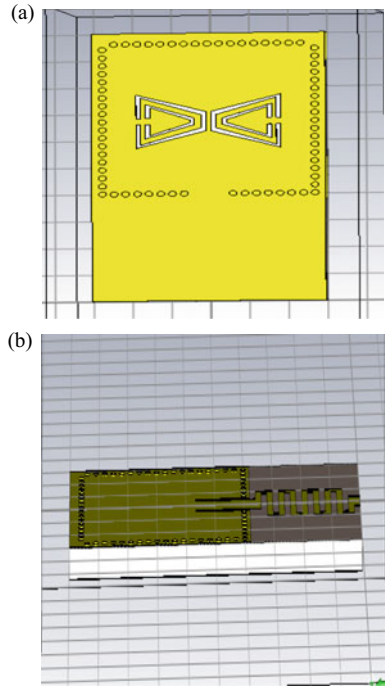
has a ground plane that is $Ls \times Ws$ in size. The meandering line radiator increases the effective current path, allowing the structure to be more compact. Equations (2) and (3) are used to estimate the substrate integrated waveguide cavity dimension in order to retain dominating frequency at 17.636 GHz, which is then optimized using CST. A 50 Ω CPW transmission line excites the SIW cavity. Impedance matching is accomplished via the inset feeding method. Figure 2 represents the proposed antenna front view and side view design.

3 Simulation Results

The reflection coefficient, radiation pattern, and gain of the planned antenna were all modeled using CST Microwave Studio. Figure 3 depicts the fluctuation of the proposed antenna's simulated reflection coefficient and gain as a function of frequency. VSWR is sensible for generated radiation across a wide frequency range with an impedance bandwidth of 233 MHz extending from 16.591 to 16.351 GHz. It is ideal for some applications because of this property. As a result, the resulting reflection coefficient suited toward the frequency 17.636 GHz, through a gain of 5.644 dB. The suggested antenna's 3D radiation pattern is depicted in Fig. 4, and it is a directional antenna with a gain > 5 dB.

(a)

(b)
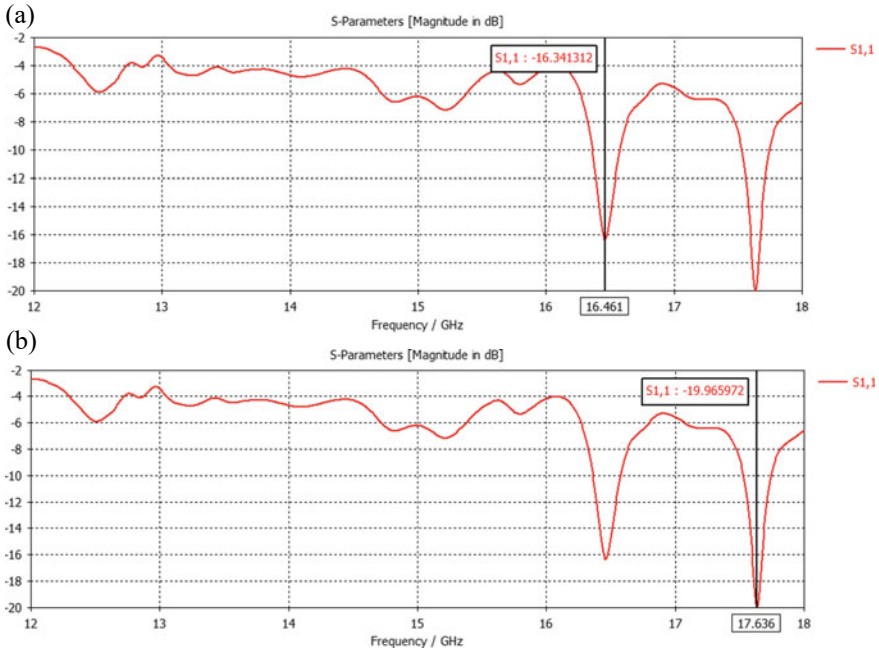
Fig. 3 Proposed antenna simulated reflection coefficient at **a** 16.461 GHz and **b** 17.636 GHz

The suggested antenna's radiation patterns were simulated at 16.461 GHz and 17.636 GHz, respectively. The value of phi = 900 has been considered at two different frequencies in order to get insight and a better knowledge of the antenna radiating behavior. In broad, radiation characteristics can be alienated into two categories: field pattern and power pattern.

In Figs. 5, 6 and 7, the antenna 3D radiation patterns at 16.461 GHz, 17.636 GHz, and at 17.636 GHz are shown. At 17.636 GHz, designed antenna gives a directivity of 10.41 dBi and gain of 5.644 dB. Figures 8 and 9 show the radiation patterns of our proposed antenna in polar representation for the frequencies 16.467 GHz and 17.646 GHz in the *H* plane and the *E* plane (*YZ* plane: Phi 90°).

4 Conclusion

The designed SIW antenna in this article presents a high gain with cavity backed triangular slots. The suggested antenna findings show gain and directivity of 3.107 dB and 8.316 dBi at 16.461 GHz and 5.644 dB gain and 10.14 dBi directivity at 17.636 GHz, respectively. Due to the presence of metalized vias, power leakage was significantly reduced. Compared to standard metallic cavities, this antenna is easier to incorporate through supplementary planar circuits and has wider operating bandwidth. The

(a)

(b)
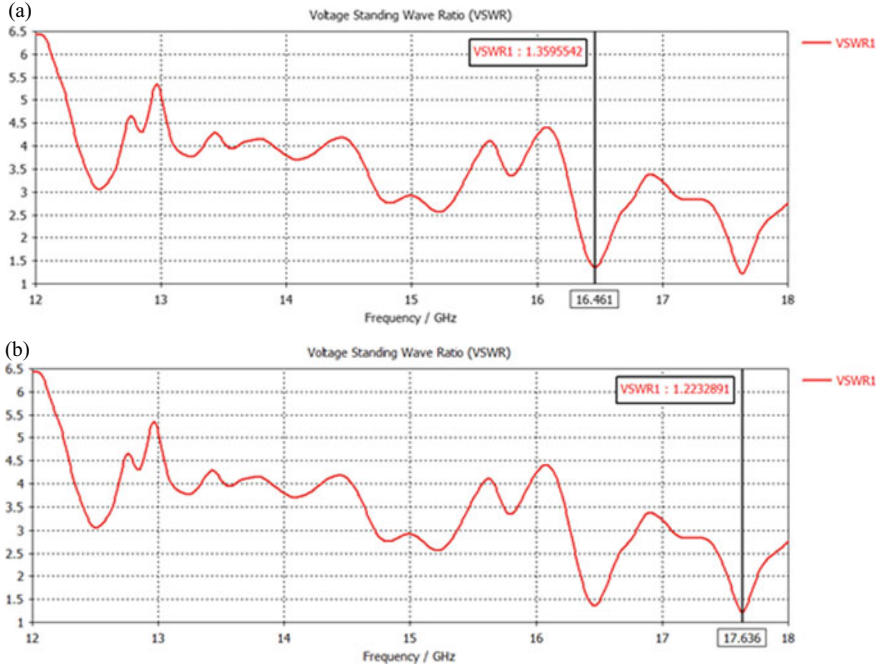
Fig. 4 Proposed antenna simulated VSWR at **a** 16.461 GHz and **b** 17.636 GHz

Type	Farfield
Approximation	enabled (kR >> 1)
Monitor	farfield (f=16.467) [1]
Component	Abs
Output	Gain
Frequency	16.467 GHz
Rad. effic.	-5.208 dB
Tot. effic.	-5.419 dB
Gain	3.107 dB
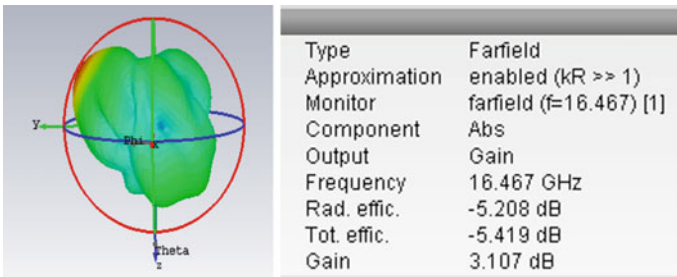
Fig. 5 Proposed antenna 3D radiation pattern at 16.461 GHz
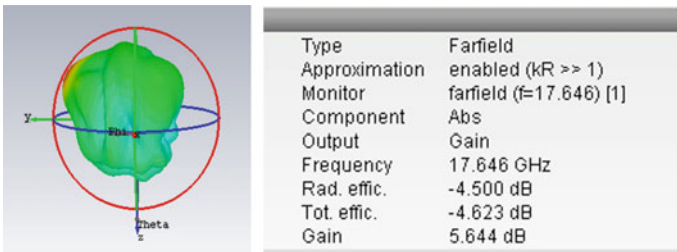
Type	Farfield
Approximation	enabled (kR >> 1)
Monitor	farfield (f=17.646) [1]
Component	Abs
Output	Gain
Frequency	17.646 GHz
Rad. effic.	-4.500 dB
Tot. effic.	-4.623 dB
Gain	5.644 dB

Fig. 6 Proposed antenna 3D radiation pattern at 17.636 GHz

Type	Farfield
Approximation	enabled (kR >> 1)
Monitor	farfield (f=17.646) [1]
Component	Abs
Output	Directivity
Frequency	17.646 GHz
Rad. effic.	-4.500 dB
Tot. effic.	-4.623 dB
Dir.	10.14 dBi

Fig. 7 Proposed antenna 3D radiation pattern at 17.636 GHz

Frequency = 16.467 GHz

Main lobe magnitude =   2.68 dB

Main lobe direction =   114.0 deg.

Angular width (3 dB) =   42.2 deg.

Side lobe level =   -7.0 dB
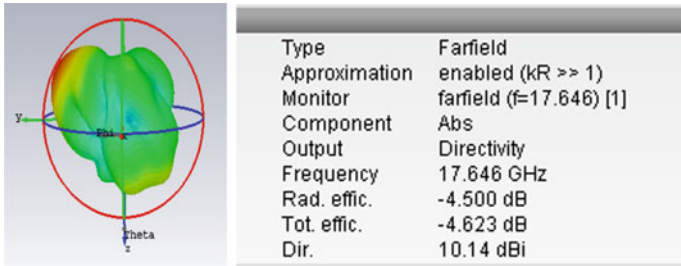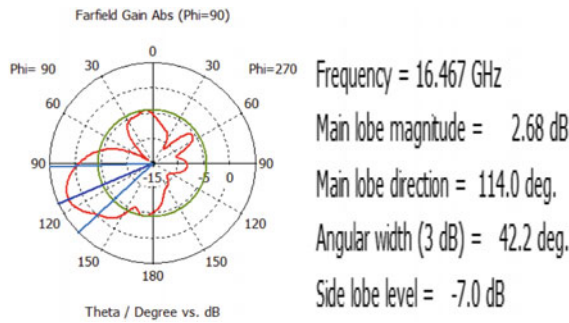
Fig. 8 Proposed antenna 2D field pattern at 16.467 GHz

Frequency = 17.646 GHz

Main lobe magnitude =   3.61 dB

Main lobe direction =   117.0 deg.

Angular width (3 dB) =   29.5 deg.

Side lobe level =   -3.7 dB
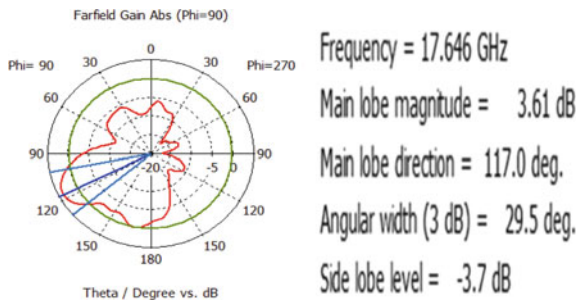
Fig. 9 Proposed antenna 2D field pattern at 17.646 GHz

proposed structure's unique properties, such as its radiation and gain distinctiveness, make it a potential contender for realistic application. Furthermore, the planned antenna component is effortlessly adaptable toward array configurations.

References

1. Saghati P, Entesari K (2017) A ultra-miniature SIW cavity backed slot antenna. IEEE Antennas Wirel Propag Lett 16:313–316
2. Tiwari B, Gupta SH (2020) Design and comparative analysis of compact flexible UWB antenna using different substrate materials for WBAN applications. Appl Phys A 126(858)
3. Tiwari B, Gupta SH, Comparative exploration of diverse substrate materials on performance of ultra wide band antenna design for on body WBAN applications. Wirel Pers Commun 1–24
4. Liu L, Wang H, Zhang Z, Li Y, Feng Z (2015) Wideband substrate integrated waveguide cavity-backed spiral-shaped patch antenna. Microw Opt Technol Lett 57:332–337
5. Zhang XH, Luo GQ, Dong LX (2013) Substrate integrated waveguide fed cavity backed slot antenna for circularly polarized application. Int J Antennas Propag 2013:6
6. Bozzi M, Georgiadis A, Wu K (2011) Review of substrate-integrated waveguide circuits and antennas. IET Microw Antennas Propag 5:909–920
7. Garg S, Kumar R (2015) Multiband microstrip patch antenna for wireless applications using metamaterial. Int J Adv Res Electron Commun Eng (IJARECE) 4(6) (2015)
8. Malisuwan S, Sivaraks J, Madan N et al (2014) Design of microstrip patch antenna for Ku-band satellite communication applications. Int J Comput Commun Eng 3(6):413
9. Vijayvergiya PL, Panigrahi RK (2017) Single-layer single-patch dual band antenna for satellite applications. IET Microw Antennas Propag 11(5):664–669

Social and Mental Well-Being-COVID-19 After Effects Survey and Data Analysis

Manasvi Narayan, Shreyash Chaudhary, and Oshin Sharma

Abstract COVID-19 pandemic affected the entire globe in 2019. This pandemic is considered as the first one with its defense more than pharmaceutical measures such as: Personal hygiene (hand sanitization, wearing masks) and social distancing. This pandemic has affected people with different factors such as: anxiety, emotions, social life, mental and physical health, and economic crisis. These factors helped this pandemic to turn up with digital solutions for its prevention and prediction estimation. These prediction techniques can analyze the previous data set of this pandemic and provide interesting insights about such a situation to occur in future along with its prevention measures. In this article, we tried to systematize various research activities-using machine learning, data science, and data visualization to extract meaningful information about COVID-19. Data collection has been done by conducting open surveys on different platforms such as: social media, university survey as well as community survey. Based on the collected data, analysis has been done on the emotional, social and mental health of people in order to provide future research directions and collective fight against such pandemics.

Keywords COVID-19 · Pandemic · Data analysis · Machine learning · Mental health · Emotional health · Social life

1 Introduction

March 11, 2020, was the day when WHO-World Health organization declared COVID-19 as pandemic. At that time, there were approximately 629,342 deaths and 13 million active cases across the world (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports accessed on September 15, 2021). The spread, if this started from Wuhan city of China, and moved toward other countries, almost all around the globe. Soon after the spread, people started following social distancing to minimize the spread. Indian Government declared nationwide

M. Narayan · S. Chaudhary · O. Sharma (✉)
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Delhi-NCR, Modinagar, Ghaziabad, Uttar Pradesh, India
e-mail: oshinsharma40@gmail.com

lockdown from March 25, 2020 to April 14, 2020. Restrictions of outdoor gatherings made people less active, less interactive, increased their screen time which further cause negative impact on their lifestyles such as: anger, frustrations, irregular sleeping patterns, obesity, and so on [1]. This homebound condition was the main reason for disappointment, insecurity, future stability, economic crisis, irritability, loneliness, and depression among people [2]. Fear of getting infected, being helpless, being sick, and not getting enough medical resources or dying became triggers for mental breakdown. It has been clear that coronavirus SARS-CoV-2 has affected everyone-mentally or physically. Here, digital technologies play a significant role and provide modern effective solutions for people in the healthcare sector. New era of machine learning, artificial intelligence, data science, and big data made remote monitoring possible. The main source for these digital technologies to provide some new insights about COVID-19 after effects is data. Many survey articles suggested the use of big data and Artificial Intelligence (AI) to detect and predict the accurate number of Covid cases [3, 4]. AI technologies provide preventive and predictive healthcare environments. It predicts the flow of virus, level of infection and estimate of healthcare facilities. Thus, AI is useful for such estimations and predictions from data generated from different resources such as: diagnosis, pharmaceuticals, medical resources, and risk score of patients [5, 6]. Moreover, with respect to healthcare, AI is the combination of both machine learning (ML) and deep learning (DL) and used for the investigation of structured data [7]. This study provides detailed insights about the actual use of AI and ML in COVID-19 pandemic. This study is about the social and mental health of students as well as working people during lockdown and after effects of lockdown on their career development like [3] have done in cross-national survey. This paper is organized in different sections followed by introduction. Second sections discuss about the different tools, techniques, and methodology adopted to conduct this research.

2 Methodology

To discuss the mental and social health of people, we have first collected the data by conducting a survey through Google Forms. We have used various ways for the distributions of this survey such as messaging service, emails, and social media, and in person [8]. Responses generated from Google form were visualized and analyzed using the concepts of Python. We have divided the survey into three different sections: (a) *social interactions during lock down*, (b) *house hold expenditure*, (c) *Covid-19 sufferings*. These three different sections provide qualitative and quantitative data from responses. Once the data has been collected, we have performed analysis to find out the hidden insights about the impact of COVID-19 on social and mental health of every individual.

Table 1 Health of participants during first and second Covid waves

Health effected during COVID-19	Counts	Percentage (%)
First wave	123	30
Second wave	196	47.6
Similar during both time	59	14.3
Not affected at all	33	8

2.1 Participants and Their Characteristics

This survey includes participants from India (95%) and three most common types of participants are: students (63%), working professionals (35%), and not employed (2%). Out of which 85% of people or their family members suffered from COVID-19. Majority participants are from State Uttar Pradesh (29.5%) and least number of participants are from Punjab (0.47%). Table 1 shows the health of participants during both First and second wave of COVID-19.

2.2 Statistical Analysis

Data preprocessing and aggregation were performed using python programming along with NumPy and pandas libraries. Statistical tests were performed on demographical data to find out the relationship between respondents and their respective responses. This test determined how two sets of pairs are statistically different from each other. Two python libraries such as Statsmodel and scipy were used to implement these statistical hypotheses [7]. Chi square test has been performed on student centric categorical data such as: feelings due to change in studies and overall frequency of negative emotions. About 95% of confidence intervals along with P value < 0.05 accept the alternative hypothesis that frequency of negative emotions in students increases when they are stressed with change in study mode. Scipy function in Python has been used to implement Pearson Chi-Square test.

2.3 Exploratory Data Analysis

Exploratory data analysis (EDA) refers to the process of finding initial investigations on data to find anomalies, detecting outliers, Univariate, and multivariate analysis of data. Here, EDA follows with data cleaning of the unstructured data which we got from the participants by removing irrelevant and incorrect information. Structural errors such as: naming conventions, incorrect capitalization, and typo errors have been fixed. For example, Timestamps were converted into date time format for better analysis of data and responses. Furthermore, student centric data and
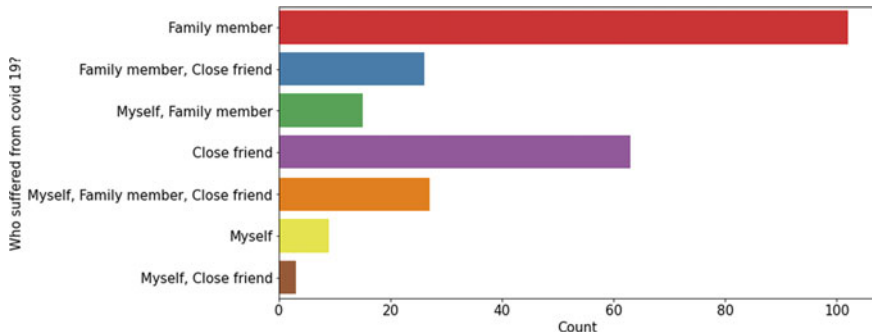
Fig. 1 Count of people suffered from COVID-19

employees related data has been cleaned by removing a few irrelevant columns such as: "*occupation affected by covid19*", "*Did you receive any incentives or financial help from your organization?*" from student centric data and "*How your study affected from online mode of teaching?*", *Did your household expenditure increases or decreases?*". Univariate and multivariate analysis has been performed on data and we got to know about every individual's mental health. After that, outliers have been detected and removed from the data such as: we got three International responses as outliers and that's why we removed country columns from our data and made our study nation centric. Figure 1 shows the interpretation of people suffered from Covid.

3   Results and Discussion

This section discusses the different analysis performed on collected dataset. We have analyzed how online mode of teaching and working affected social interaction, emotional well-being and mental health of individuals.

3.1   *Assessment of Social Interaction During Lockdown*

Stress and anxiety were two main consequences of social distancing and home arrest situations during Covid lockdown. To assess the effects of lockdown and online mode of teaching we asked the students 4–5 different questions related to the social interactions of students such as: level of interaction from 0 to 5, mode of communication and relationship with friends and family, and the effect of changed mode of interaction. Most of the students ranked their interaction with friends declined and the average rating is 2.5. However, the majority of them mentioned improved interaction
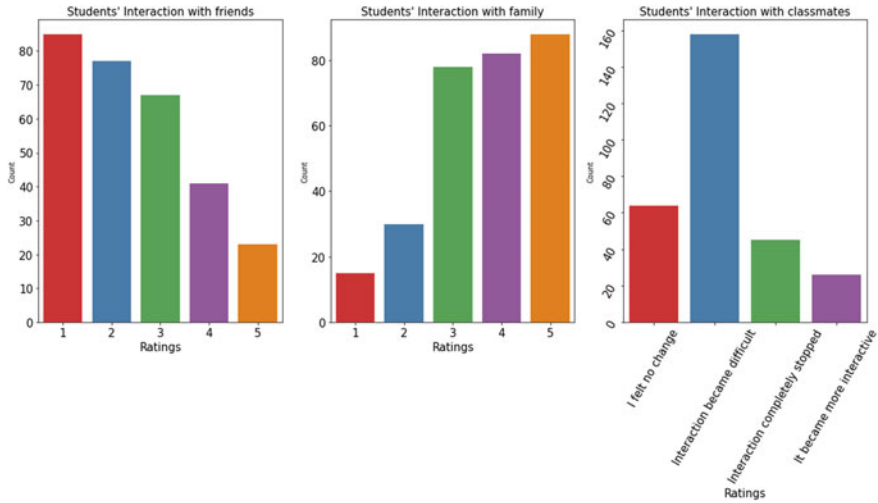
Fig. 2 Students' interaction with their friends, family, and classmates

with family with average rating 3.8 and interaction with classmates became difficult or declined during Covid lockdown (Fig. 2).

To discuss the impact of work from home or closed business on family relationships we have also taken responses from participant's and found that phone calls were the main source of communication during lockdown. Social media played an important role for the online mode of connectivity with friends and family. Average rating for family interaction is 3.75. Few people mentioned that they felt no change in interaction during both the first and second wave of Covid. Figure 3 show that the people have improved the connection with the family. They came close, as due to lockdown, they had to stay at home, do the work and spend time with family. It also shows that colleagues have faced many difficulties in communication. About 46% of the people said that they felt difficulty in interaction with colleagues. Also, about 30% of people said that they felt no change in communication and their communication was smooth.
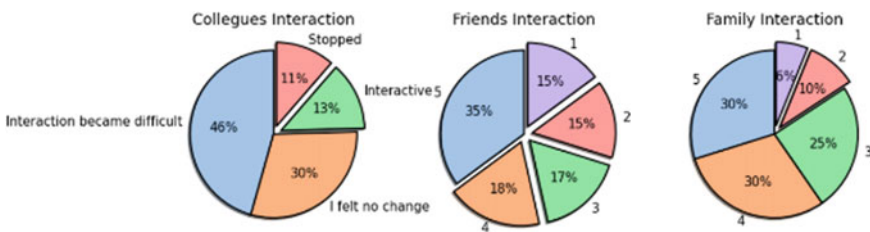
Fig. 3 Social interaction with colleagues, friends, and family
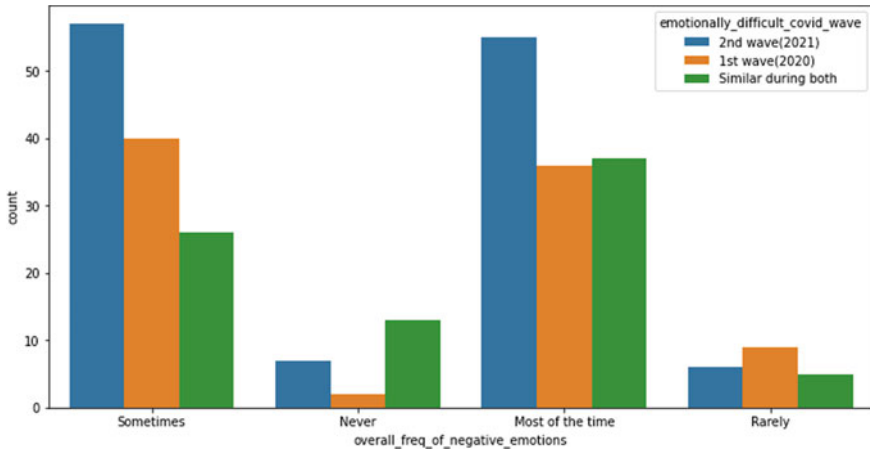
Fig. 4 Overall frequency of negative emotions during first and second Covid waves

3.2 Assessment of Emotional Well-Being During Home Arrest Situation

To assess the emotional well-being of people we asked participants six different questions related to their emotions such as: level of loneliness, level of frustration being confined at home, number of times being a part of family argument, level of feeling depressed: (never-rarely-most of the times-sometimes), and effects of changed mode of teaching: (not relevant, relaxed, normal, and stressed). Depending upon the responses of people, a new feature "feelings of negative emotions" has been created to assess the emotional well-being of students and for which we have set the priority mode for the answers "Most of the times > sometimes > rarely > never ". Label encoding has been done on following columns of data: to encode the categorical variables such as:—Never: 0—Rarely: 1—Sometimes: 2—Most of times: 3, answers:—yes: 2, No: 1, May be: 0. We have assessed emotional well-being using negative emotions and to calculate the frequency of negative emotions of every individual we have used decision trees (Fig. 4).

People who had to make extra healthcare expenditure and either themselves or some family member or friend were suffering from severe Covid (rating ≥ 3) most often felt downhearted.

3.3 Assessment of Mental Health of People During Covid Lockdown

To assess the mental health of every individual we asked the six questions and analyzed their responses. Factors that cause mental and emotional health are: change

in work (work from home), families and friends suffering from Covid, increase in house expenditure, expenses during Covid illness. About 37% of people said that they felt frustrated most of the time and 49% people felt depressed sometimes. According to the decision tree which we have created in Sect. 3.2 we got to know about the importance of different features and how these features are creating negative emotions in every individual. Table 2 shows the individual count for different feelings and emotions. We observed that people from the private or corporate sector have negative emotions most of the time during Covid lockdown. Moreover, people's mental health has also been assessed by considering the level of severity that the patient himself/herself or family members or close friends felt from (mild-severe—very severe) during Covid. It has been observed during analysis that the mental health of the majority of individuals affected when their family members were in very severe condition rather than themselves or their close friend were in severe or very severe condition (Fig. 6). 53% students and 57% working professionals were stressed when their family members were suffering from Covid. Figure 5 shows the importance of features for analyzing the mental health of people during lockdown.

Thus, we have analyzed that people were not worried about their own lives as much as they were worried or stressed about their loved ones. We have also analyzed that mental health of people also varies with respect to their occupation. Figure 7 shows how occupation affects each and everyone's emotions.

Table 2 Individual count for feelings by an individual

Feelings	Sometimes (count)	Most of times (count)	Rarely (count)	Never (count)
Anxious	51	37	15	11
Argumentative	56	10	20	28
Depressed	56	17	25	16
Frustrated	42	42	15	15
Lonely	56	23	13	22
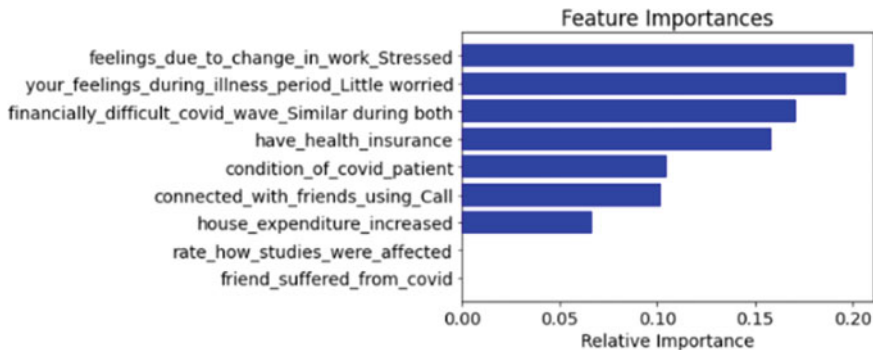
Fig. 5 Feature importance for analysis of mental health
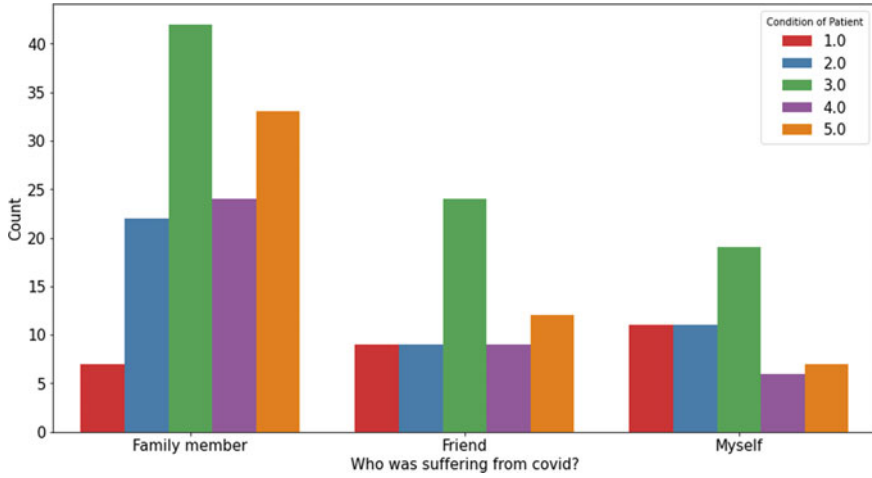
Fig. 6 Condition of patient of Covid
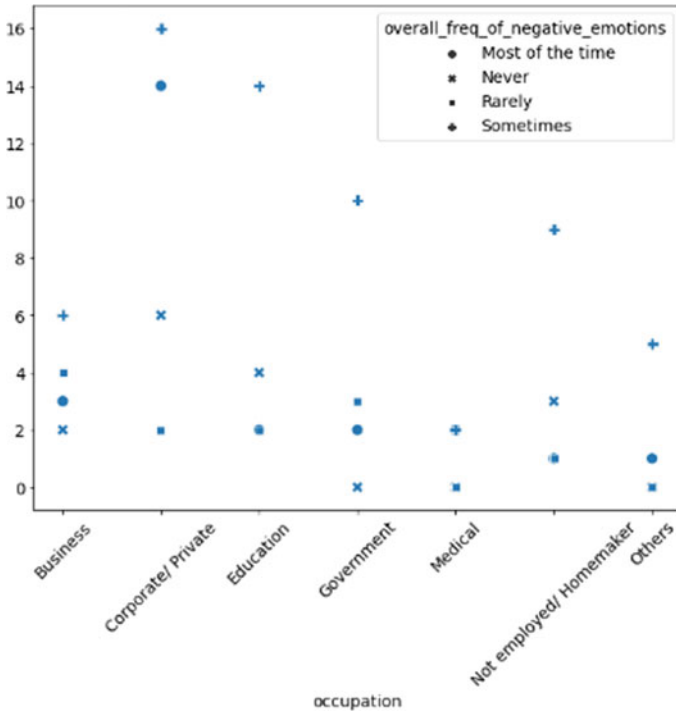
Fig. 7 How emotions affected with respect to occupation
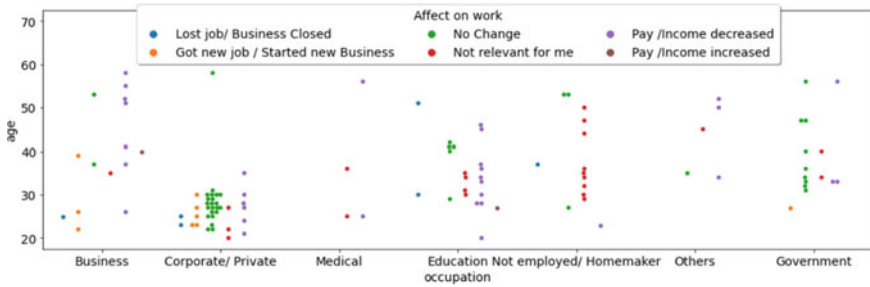
Fig. 8 Financial after-effects of Covid 19

3.4 Assessment of Economic Condition of People

The impact of this pandemic on the Indian economy was also very disruptive. We have conducted this survey to understand and analyze the financial condition of working professionals. We divided this analysis into four groups: Business, Corporate, Education Sector, and Government Jobs. About 50% of small-scale business people mentioned their change in income which further depends upon type of business. Sales in this period have increased for very few businesses. In the Corporate sector, more than 50% people said that they got new jobs, which shows that there were many openings in the different companies for different positions and nearly 48% people said that they felt no change. In the education sector, 50% of people said that their pay was increased during lockdown. That means the education sector was growing and the people got high salaries in this field. People working in the government sector said that they felt no change financially during the lockdown period. Figure 8 shows how Covid 19 affects people financially.

3.5 Future Outlook Regarding Covid and Other Pandemics

The arrival of this highly infectious variant became the reason for anxiety and panic attacks in everyone's home. Thus, we have conducted this survey to gain the insights for future outlook of such kinds of pandemics. We conducted a Chi squared test to find out what features were statistically dependent on the future outlook variables. Mental and emotional health of every individual was the deciding factor on their thoughts about the future outlook. Those who least frequently felt depressed, frustrated, and stressed were the ones who were most confident in themselves and the government to handle such situations in the future. Figure 9 shows the heat map of feature dependency on how people are ready to handle such pandemics in the future.
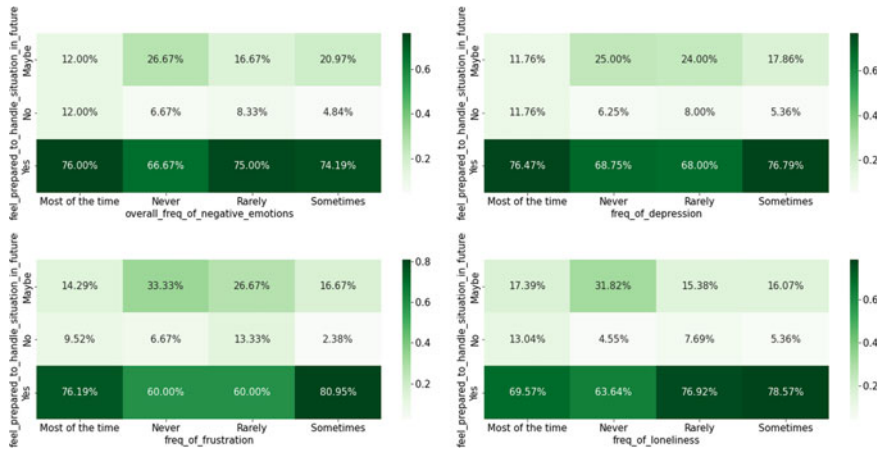
Fig. 9 Heat map of feature dependency on how people are ready to handle such pandemics in future

4 Conclusion

COVID-19 lockdown has made significant disruptions in our day today's life and activities. This study assessed the mental, social and emotional health of people suffering from Covid or after affects of COVID-19. We analyzed that everyone got more involved into their families as they were staying at home, working from home and studying from home therefore, their social interaction decreased and family interaction increased. Furthermore, we investigated different features that are generating the feelings of negative emotions which further cause mental and emotional sickness. COVID-19 financial after affects has also been analyzed which discuss how different occupations get affected during nationwide lockdown situation. Future outlook of this study tells that people are ready to handle such situations in the future as they found different ways to increase their social interaction, to deal with negative emotions, to improve mental health as well. In future we would like to work on the detailed analysis of different variants of COVID-19 and how they really affect human metabolism as well as analysis of medical side effects of COVID-19's variants.

References

1. Liu JJ, Bao Y, Huang X, Shi J, Lu L (2020) Mental health considerations for children quarantined because of COVID-19. Lancet Child Adolesc Health 4:347–349
2. Horita R, Nishio A, Yamamoto M (2020) The effect of remote learning on the mental health of first year university students in Japan. Psychiatry Res 295:113561
3. Agbehadji IE, Awuzie BO, Ngowi AB, Millham RC (2020) Review of big data analytics, artificial intelligence and nature inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing. Int J Environ Res Public Health 17:5330

4. Vaishya R, Javaid M, Khan IH, Haleem A (2020) Artificial intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab Syndr Clin Res 14:337–339

5. Khan ZF, Alotaibi SR (2020) Applications of artificial intelligence and big data analytics in m-health: a healthcare system perspective. J Health Eng 1–15

6. Mervin JT, Vishnu L, Ajith KB, Muhammad RVP, Alosh J, Arun KR (2021) Can technological advancement help to alleviate Covid 19 pandemic? a review. J Biomed Inf 117

7. Saha J, Barman B, Chouhan P (2020) Lockdown for COVID-19 and its impact on community mobility in India: an analysis of the COVID-19 community mobility reports, 2020. Child Youth Serv Rev116

8. Kapasia N, Paul P, Roy A (2020) Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal, India. Child Youth Serv Rev 116

9. Daniela P, Andre G, Francesco R, Jorge C, Emanuele DF, Sofia GC, Emilio Z (2021) Behaviours and attitudes in response to the Covid-19 pandemic: insights from a cross national facebook survey. EPJ Data Science. 10:17

10. International Labour Organization. Impact of covid-19 crisis on loss of job. Available online: https://www.ilo.org/global/topics/domestic-workers/publications/factsheets/WCMS_7 47961/lang--en/index.htm. Accessed on 21 Sept 2021

11. World Health Organization. Coronavirus disease (COVID-2019) situation reports. Available online: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-rep orts. Accessed on 15 Sept 2021

Covid Patient Monitoring System for Self-quarantine Using Cloud Server Based IoT Approach

Mettu Jhansi Lakshmi, Gude Usha Rani, and Baireddy Srinivas Reddy

Abstract A pulse oximeter is a piece of medical equipment that analyses the quantity of oxygen concentration in a person's bloodstream, i.e., what proportion of the oxygen-carrying molecules known as hemoglobin) simply transport oxygen around the body. Pulse oximetry is based on the assumption that two wavelengths may be utilized to make arterial blood oxygen choices, assuming that the observations are done on the pulsatile component of the signal. Traditional pulse oximeters employ two leads at varied wavelengths and a phototransistor to estimate blood oxygen content without being intrusive, defined as the variation in coefficient of reflection between hemoglobin and deoxyhemoglobinin. Here, we proposed a self-quarantine system to monitor the condition of the patient remotely using cloud technology. The system monitors the temperature and SpO_2 level continuously and uploads it to a private server in certain interval of time. So, the doctor or caretaker can view the condition of the patient remotely.

Keywords Cloud server · Covid monitoring system · SpO_2 · Oxygen saturation · Self-quarantine

1 Introduction

Like most viruses, Corona is transmitted when an infected person coughs or sneezes on someone else. You can also become infected by breathing in Corona-infected sweat particles.

M. Jhansi Lakshmi (✉) · G. Usha Rani · B. Srinivas Reddy
Department of Information Technology, CMR Engineering College, Hyderabad, Telangana, India
e-mail: jhansi2023@gmail.com

COVID-19 patient monitoring system is a fully-automated, on-demand oxygen delivery and treatment system designed to quickly provide life-saving treatments to patients in emergency situations developed and built by team members at Aethlon Medical, COVID-19 uses a unique combination of advanced artificial intelligence and 3D printing to deliver targeted doses of oxygen to patients with few or no symptoms of hypoxia. It is capable of delivering high-quality oxygen to patients in a fraction of the time required by traditional treatments. It is able to provide a continuous supply of oxygen to patients without interruption, reducing the burden on emergency responders and hospitals.

Oxygen is the most common element in the universe, making up more than twenty-eight percent of the entire mass of the planet earth. It's also one of the most abundant elements in the human body, making up roughly twenty percent of your mass. The most common type of oxygen we encounter is molecular oxygen-this is the stuff we breathe on a daily basis. But, molecular oxygen is only one type of oxygen; there's also a different type of oxygen that's only found inside living organisms, and another type of oxygen that's only found in the environment. So, here the proposed system uses WSN oxygen saturation monitoring of the patient to monitoring. Pulse oximeters have become an essential approach for healthcare management, ranging from emergency departments and clinic hospital beds to sleep disruption diagnosis in the clinical household. They use photoplethysmography to determine the quantity of oxygen in the blood stream non-invasively (SpO_2).

2 Literature Review

In this study [1], we present a cheap IoT-based system aimed for increasing COVID-19 indoor safety by encompassing numerous key factors such as contactless temperature monitoring, mask detection, and social distancing check. The contactless temperature sensing subsystem is powered by an Arduino Uno and an infrared sensor or a thermal camera, while mask recognition and social distance checks are handled by computer vision algorithms on a camera-equipped Raspberry Pi.

This study [2] discusses how the Internet of Things (IoT) may be integrated into the epidemic prevention and control system. We show a potential fog-cloud combined IoT platform that can be used in the systematic and intelligent COVID-19 prevention and control, which includes five interventions: COVID-19 symptom diagnosis, quarantine monitoring, contact tracing and social distancing, COVID-19 outbreak forecasting, and SARS-CoV-2 mutation tracking [3].

Based on this backdrop, the authors previously presented an IoT-based healthcare infrastructure to allow remote monitoring for patients in critical situations. As a result, the purpose of this article is to broaden the platform by including wearable and inconspicuous sensors to monitor patients with coronavirus illness. Furthermore, we describe a real-world use of our technique in a COVID-19 critical care unit in Brazil [4].

For classifying COVID-19 patients, researchers presented a variety of machine learning and smart IoT-based methods. Artificial neural networks (ANNs) are commonly utilized in a variety of applications, including healthcare systems, and are inspired by the biological idea of neurons. In the decision-making process for handling healthcare information, the ANN scheme provides a suitable option [5].

This study [6] proposes a new approach of detecting COVID-19 fever symptoms based on IoT cloud services to address the longer time delay of monitoring packed customers who enter public or private agencies, which can lead to a risky field for disease propagation. A realistic experiment is used to design an autonomously checking procedure.

3 Proposed Methodology

Pulse oximetry is typical medical measuring equipment that monitors the concentration of oxygen saturation in our circulation and can detect minute fluctuations in oxygen. It is a non-invasive and painless test. It is vital in the current COVID-19 scenario to monitor patient the oxygen saturation of several patients who required without getting into close contact with patients.

In this paper, we presented a self-quarantine system that uses cloud technologies to remotely monitor the patient's status (Fig. 1). The system continually checks the temperature and SpO_2 level and uploads it to a private server at regular intervals. As a result, the doctor or caregiver can monitor the patient's status from afar.

Pulse oximetry is a noninvasive medical test that uses light to measure the amount of oxygen in the blood. It's used to monitor the oxygen levels of people who are unable to breathe on their own, such as those with breathing disorders, lung diseases,
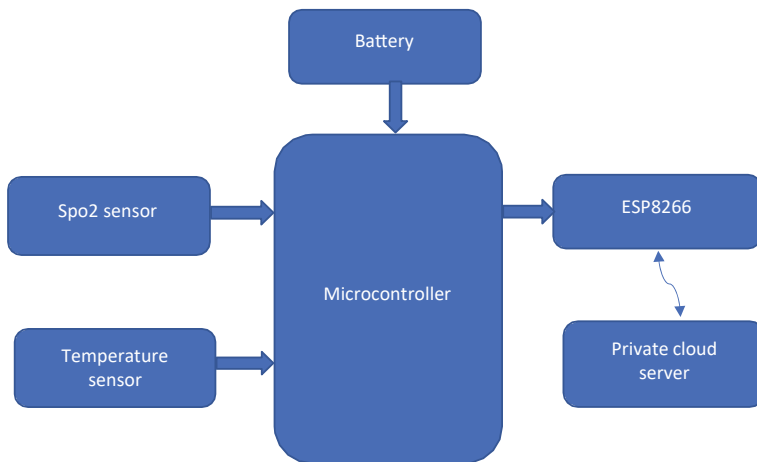
Fig. 1 Proposed model
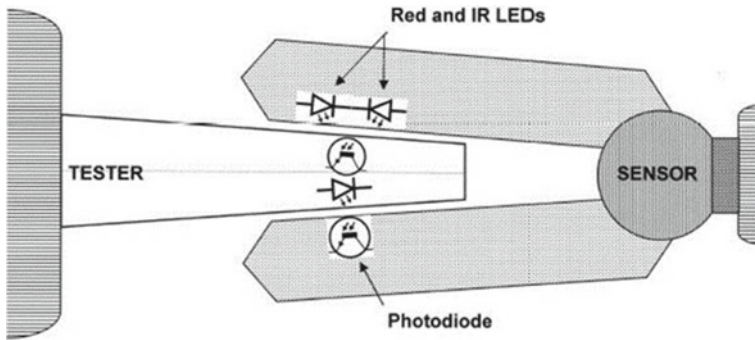
Fig. 2 SpO$_2$ working

or cardiac arrest. The test is painless, requires only a few seconds of exposure to bright light, and is frequently given without the patient's knowledge. A pulse oximetry test can be performed at home using a device called a pulse oximeter.

However, when used on patients with certain medical conditions, such as hypoxemia, pulse oximetry can provide vital information that can help doctors better understand and manage their condition. Because of their ability to provide real-time oxygen level data, pulse oximeters have also become an important tool on the battlefield and in first aid kits.

The most common method of pulse oximetry is through the use of a small probe called a sensor, which is placed on a patient's finger, toe, or earlobe. When the sensor is placed on a patient's body, it emits small amounts of red and infrared light. The amount of red and infrared light that is absorbed by the tissues in the body is then measured by a device called a pulse oximeter (Fig. 2).

4 Results and Discussion

The world of healthcare is changing. Patients now have more control over their healthcare than ever before, and this means that they can also make decisions about their care. One of the biggest healthcare decisions that patients can make is where they want to be treated. Some patients want to be treated at home, while others prefer to be treated in a hospital or care facility.

Covid health is a telemedicine platform that allows patients to access healthcare from the comfort of their homes. One of the ways we help our patients is by providing them with access to our MV telemedicine platform, which allows them to video chat with our doctors and specialists who can diagnose them and prescribe them medications. One of the ways we help our patients is by providing them with a temperature sensor that can be used to monitor their body temperature, especially in cases of suspected infection. This allows our doctors to proactively prescribe antibiotics and other medications, which is especially important for patients who

Fig. 3 Cloud result
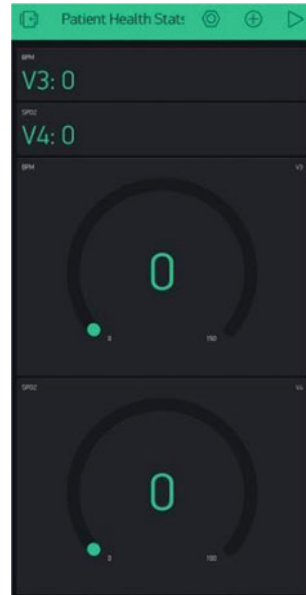
don't have access to urgent care or don't want to go to the doctor for every little thing.

As demonstrated in Fig. 3, we may use the blynk app to capture individual sensor data and input it into the cloud server. The hospital management server database can monitor this data. These are all IoT-connected and can be an effective means of gathering data from patients.

This research resulted in a graph for that particular design of the SpO_2 transmitters and display. This is referred to as an *R*-curve.

An *R*-curve is defined as the correlation between the fundamental ratio of red and near-infrared light versus the observed oxygen levels acquired during human testing, as shown in Fig. 4. The R-curve is then utilized in the framework for a specific item and for SpO_2 testers.

The temperature of most devices electronic or otherwise is measured in degrees Celsius. However, many electronic devices have a temperature sensor on them which is measured in degrees Kelvin. The difference between the two is simple enough: one is absolute, the other is relative. In the Fahrenheit scale, zero degrees is the freezing point of water, 100° is the boiling point of water. Figure 5 shows the graph representation of temperature and voltage.
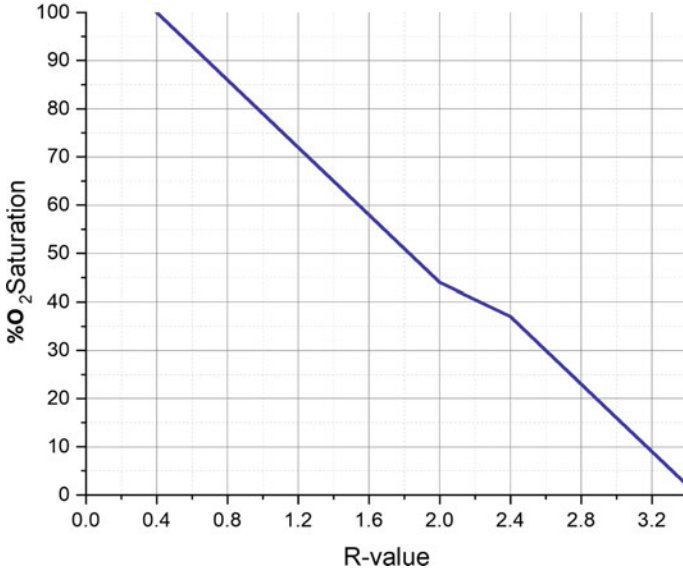
Fig. 4 SpO$_2$ saturation
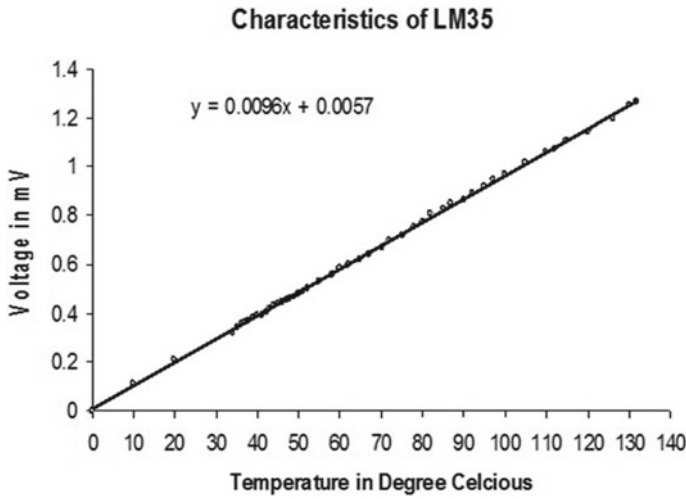
Fig. 5 Temperature versus voltage

5 Conclusion

Self-quarantine is much needed for nowadays to reduce the transmission of the virus. Pulse oximetry is predicated on the notion that two wavelengths can be used

to determine arterial blood oxygen levels, given that the measurements are made on the pulsatile component of the signal. The established IoT-Enabled Pulse Oximeter has proven a moderate success when compared to conventional Pulse Oximetry and Pulse-Rate equipment. As a result, we have thoroughly validated and tested our objective of making the gadget functional and small. Even faraway physicians can assess a patient's health by checking the outcome from the cloud using an Internet connection. We introduced a self-quarantine system that employs cloud technology to remotely monitor the patient's state in this research. At regular intervals, the system monitors the temperature and SpO_2 level and transmits it to a private server. As a consequence, the doctor or caregiver can keep an eye on the patient's condition from afar.

References

1. Petrović N, Kocić Đ (2020) IoT-based system for covid-19 indoor safety monitoring. Preprint. IcETRAN 2020, pp 1–6
2. Dong Y, Yao Y-D (2021) IoT platform for COVID-19 prevention and control: a survey. IEEE Access 9:49929–49941
3. Jung Y, Agulto R (2021) A public platform for virtual IoT-based monitoring and tracking of COVID-19. Electronics 10(1):12
4. de Morais Barroca Filho I, Aquino G, Malaquias RS, Girão G, Melo SRM (2021) An IoT-based healthcare platform for patients in ICU beds during the COVID-19 outbreak. IEEE Access 9:27262–27277
5. Rathee G, Garg S, Kaddoum G, Wu Y, Jayakody DNK, Alamri A (2021) ANN assisted-IoT enabled COVID-19 patient monitoring. IEEE Access 9: 42483–42492
6. Kamal M, Aljohani A, Alanazi E (2020) IoT meets COVID-19: status, challenges, and opportunities. arXiv preprint arXiv:2007.12268

Wood Images Classification Based on Various Types of K-NN Classifier

Madhuri R. Kagale and Parshuram M. Kamble

Abstract A classification model for the automatic classification of wood images using the computer vision approach is presented in this proposed study. A very little work has been done on automatic wood classification using various types of K-nearest neighbor(K-NN) classifier algorithms. The aim of this research is to compute textural features using gray-level co-occurrence matrix (GLCM) from the wood images and classify them by implementing different types of K-NN classifier algorithms. Results obtained by experiments show that a Weighted K-NN classifier computes better classification accuracy among other K-NN classifiers used in this study. The methodology proposed in this study is effective for automatic wood classification, which has several industrial applications.

Keywords Wood images · Classification · Extracted features · Weighted K-NN algorithm

1 Introduction

Texture classification is an effective method to solve numerous problems in texture analysis study. Most of the classification methods were implemented on grayscale textures. Texture plays an important role to classify texture images from the real world. A wide variety of texture classification methods are used in numerous computer vision applications like object recognition, segmentation, face recognition, and wood identification in the industry.

Wood plays an essential role in various industrial applications over the past few decades. Wood is an extensively used material in construction, furniture, interiors, and shipbuilding. Hence, accurate classification of wood is necessary to identify wood for the saw-milling process. The selection of the wood is determined by its quality

M. R. Kagale (✉) · P. M. Kamble
Department of Computer Science Central University of Karnataka, Kalaburagi 585367, India
e-mail: madhurikagale@cuk.ac.in

P. M. Kamble
e-mail: parshuramkamble@cuk.ac.in

and price. While choosing the wood for flooring, stringent selection or classification procedures are required to obtain satisfactory results. The fiber type and color tone are used to grade wood for selecting beautiful, defect-free, and uniform surfaces. To have a global presence in the competitive market, the wood industry requires a fast and accurate method to identify wood slabs. In order to choose the appropriate quality of wood material, producers are in search of the best approach for wood identification. In industry, wood classification is an essential method to identify suitable wood material based on color wood images. Earlier, the identification of wood quality was performed by manual procedure, which required trained staffing. The manual identification of wood slabs takes a longer time and laborious work. This can lead to significant variations causing sales returns and severe economic losses. This serves as a huge motivation for the development of an automated wood classification approach. This paper proposes a methodology for wood identification based on automated classification in view of color texture analysis of wood images. Various types of K-nearest neighbor (K-NN) classifier algorithms are examined for the classification that is applied to extracted texture features from the wood images of the Parquet database.

The highlights of this paper are categorized into the following sections. A related research is briefly viewed in Sect. 2. The methodology used for experimentation for the classification of wood images is provided in Sect. 3. Section 4 presents the experimental setup and results. In Sect. 5, a conclusion is made based on the present study.

2 Related Research

The wood industry has started implementing computer vision methods for achieving standard results in the last two decades. In the literature, it has been clearly mentioned that wood quality inspection is categorized into two types of problems: one is detection and classification of surface defects, and the other is sorting and classifying products that have very much similar appearance [4]. The first problem of detecting and measuring defects is referred to as grading. Many researchers have proposed various texture analysis methods to deal with the grading problem [5, 7, 11, 18].

The second is referred to as sorting and color classification of wood images. The industry is concerned with developing automatic computer vision algorithms to identify wood images to overcome the limitations of manual quality control procedures. Different methods are reported in the literature to deal with sorting and classification of wood materials [13, 14, 17].

In the present study, the problem of sorting also called as classification of color Parquet hardwood images is focused. A wide variety of approaches have published for the sorting and classification of wood using its types, quality, and defects which are very useful for the wood industry. Hence, it is required to obtain quick and quality results by implementing various classifiers [2, 6] for wood type classification.

Bianconi et al. [3] presented a technique for sorting Parquet hardwood images with the help of statistical features calculated for different color spaces. Bombardier et al. [4] employed a fuzzy classifier, wherein mean and homogeneity features are extracted from HSV and CIE Lab colored wood images. The anisotropic diffusion approach is implemented for classifying wood images in various color spaces [9]. More recently, Porebski et al. [15] used a strategy of combining LBP bin and histogram selection to classify Parquet wood images effectively. An up-to-date survey on wood recognition and quality imaging inspection system was published [12].

3 Methodology

This section consists of a wood image database, preprocessing, feature extraction using statistical methods, and K-NN classifier for wood image classification. Experiments performed on Parquet wood image database. We applied preprocessing techniques to make images noise free. After this stage, we extracted Haralick gray-level co-occurrence matrix (GLCM) features, and finally, classification has done based on K-NN classifier.

3.1 Wood Image Database

The Parquet wood image database [3] contains 14 classes of commercial types of hardwood for flooring and cladding like IRK, OAK, and TEK. Each class image is of a different type, treatment, and finish. There are two to four tones of images in each class. These images, due to their similar appearance, are difficult to differentiate with a manual procedure. There are six to eight samples for each class tone, which are provided in 24-bit color bitmap format for varying resolutions. Since each image is of variable size, the classification and identification of the correct image is a challenging task.

3.2 Preprocessing

In the database, wood images are in heterogeneous dimension and color which has to convert in uniform dimension and color before feature extraction. We applied cubic [1] size normalization method to normalize image dimension into 400×300 pixel. Gray-level transformation [16] method used to transform color image into grayscale is shown in Fig. 1.

Fig. 1 Preprocessing of original wood image: figure **a** is in RGB color, and figure **b** is grayscale image

(a) Original Image   (b) Gray scale Image
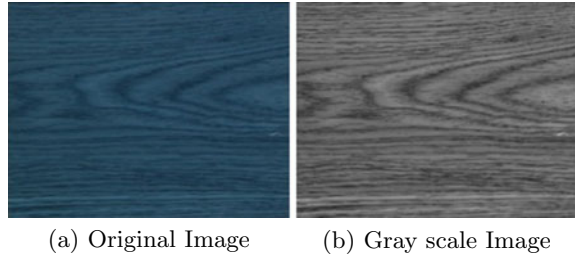
3.3 Feature Extraction

In the proposed work, we have used first- and second-order statistics to extract features from gray-level co-occurrence matrix (GLCM). A total of fifteen texture features are extracted for the study. GLCM is an extensively used statistical method to extract textural features from the images. These extracted features are considered second-order statistical properties. Out of fifteen features, there are thirteen Haralick features [8], and the remaining two are mean and standard deviation. The computational formulas for mean and standard deviation are as follows:

$$\text{mean} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

$$\text{std} = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \text{mean})}{n-1}} \tag{2}$$

3.4 Classification of Wood Images

The classification of wood images is performed using the K-nearest neighbor (K-NN) classifier. K-NN classifier mainly used in supervised learning can be implemented for different distance metrics and the number of nearest neighbors means K value [10]. Most of the researchers used K-NN classifier, but very few studies have used the various types of K-NN classifiers to classify wood images. Due to this reason, the present research is focused on applying the different types of K-NN classifiers for the automatic classification of wood images which will be helpful for the wood industry. The four different types of K-NN classifiers for the various distance metrics and the number of neighbors are used. A list of K-NN classifiers used in the study is Medium, Cosine, Cubic, and Weighted K-NN algorithms. Figure 2 shows a schematic diagram of classification for wood images.
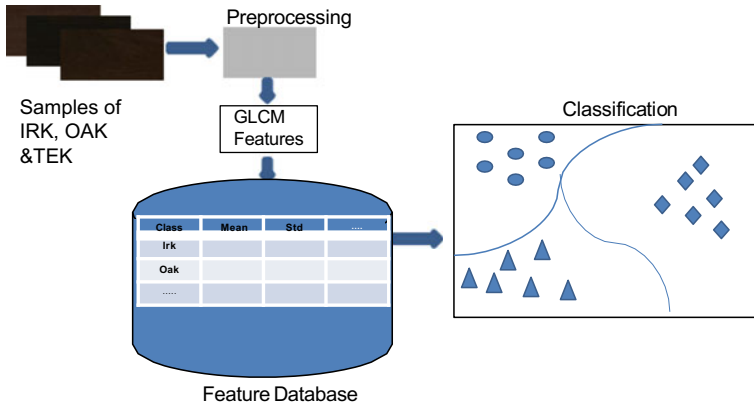
Fig. 2 Schematic diagram of classification for wood images

4 Experimental Setup and Results

The experimental tests are performed on varying size color wood images from the Parquet wood image database [3]. The database contains 14 types of wood images that are categorized into three classes, namely IRK, OAK, and TEK. The classification of texture requires two stages: training and testing. 70% randomly selected samples from each class are used for the training stage, and the remaining 30% are used for testing. And the prediction is made based on a fivefold cross-validation technique. Experimental results for various K-NN classifiers used in this study have different distance metrics and the number of neighbors (K value). The details are mentioned in Table 1. The classification accuracy obtained using Medium K-NN classifier is 93.1%. The classification accuracy achieved by using Cosine K-NN classifier is 91.7%. The Cubic K-NN classifier yields 91.7 % classification accuracy, the same as Cosine K-NN. But Cubic K-NN uses the Minkowski distance metric. Weighted K-NN uses Chebychev distance metric, squared inverse distance weight, and K value is 10. The classification accuracy using the Weighted K-NN classifier achieved 98.6 %, and its confusion matrix is given in Fig. 3.

Further, it is observed from Table 1 that the Cosine and Cubic K-NN classifier yield the same classification accuracy of 91.7% for different distance metrics. This

Table 1 Three-grade wood image classification based on K-NN classifier

Classifier name	Distance measure	No. of neighbor	Accuracy in %
Medium	Euclidean	10	93.1
Cosine	Cosine	10	91.7
Cubic	Minkowski	10	91.7
Weighted	Chebychev	10	98.6

Fig. 3 Confusion matrix for Weighted K-NN classifier
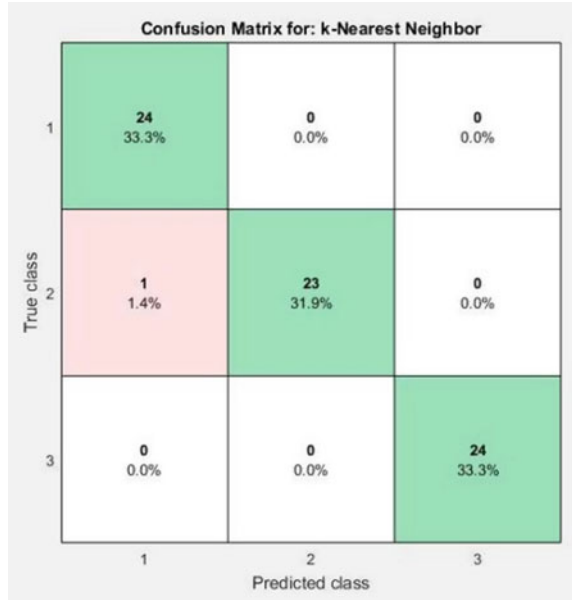
classification accuracy can be improved by a Medium and Weighted K-NN classifier algorithm. The classification accuracy obtained by the Weighted K-NN classifier is the highest among all other K-NN classifier algorithms. The Weighted K-NN classifier was found to give superior results, making it suitable for wood classification.

5 Conclusion

For the wood images classification, images from the Parquet database are categorized into three different classes: IRK, OAK, and TEK. Samples from these classes are pre-processed, and texture features are obtained using gray-level co-occurrence matrix (GLCM). Further, various types of K-nearest neighbor (K-NN) classifiers such as Modified, Cubic, Cosine, and Weighted K-NN algorithms are used to classify wood images. These classifiers are applied for different distance metrics, distance weight, and the number of neighbors, which lead to enhance classification accuracy. Experimental results show that Weighted K-NN performance is superior to other K-NN classifiers. The classification accuracy validates that the proposed method is more efficient in automatic wood classification, which can be effectively used in industrial wood applications. Further, this work can be extended to examine the other supervised learning algorithms for wood images classification.

References

1. Abu-Mostafa YS, Psaltis D (1985) Image normalization by complex moments. IEEE Trans Pattern Anal Mach Intell 1:46–55
2. Bello-Cerezo R, Bianconi F, Di Maria F, Napoletano P, Smeraldi F (2019) Comparative evaluation of hand-crafted image descriptors vs. off-the-shelf CNN-based features for colour texture classification under ideal and realistic conditions. Appl Sci 9(4):738
3. Bianconi F, Fernández A, González E, Saetta SA (2013) Performance analysis of colour descriptors for parquet sorting. Exp Syst Appl 40(5):1636–1644
4. Bombardier V, Schmitt E (2010) Fuzzy rule classifier: capability for generalization in wood color recognition. Eng Appl Artif Intell 23(6):978–988
5. Czimmermann T, Ciuti G, Milazzo M, Chiurazzi M, Roccella S, Oddo CM, Dario P (2020) Visual-based defect detection and classification approaches for industrial applications—a survey. Sensors 20(5):1459
6. Darmawan E, Novantara P, Suwarto GP, Andriyat R, Nurhayati Y (2021) The implementation of *k*-means algorithm to determine the quality of teak wood in image based on the texture. J Phys Conf Ser 1933:012003
7. Gu IYH, Andersson H, Vicen R (2010) Wood defect classification based on image analysis and support vector machines. Wood Sci Technol 44(4):693–704
8. Haralick RM, Shanmugam K, Dinstein IH (1973) Textural features for image classification. IEEE Trans Syst Man Cybernet 6:610–621
9. Hiremath PS, Bhusnurmath RA (2017) Multiresolution LDBP descriptors for texture classification using anisotropic diffusion with an application to wood texture analysis. Pattern Recogn Lett 89:8–17
10. Hwang WJ, Wen KW (1998) Fast *k*NN classification algorithm based on partial distance search. Electron Lett 34(21):2062–2063
11. Jabo S (2011) Machine vision for wood defect detection and classification. Master's thesis
12. Kryl M, Danys L, Jaros R, Martinek R, Kodytek P, Bilik P (2020) Wood recognition and quality imaging inspection systems. J Sens
13. Liu S, Jiang W, Wu L, Wen H, Liu M, Wang Y (2020) Real-time classification of rubber wood boards using an SSR-based CNN. IEEE Trans Instrum Meas 69(11):8725–8734
14. Madyan OA, Wang Y, Corker J, Zhou Y, Du G, Fan M (2020) Classification of wood fibre geometry and its behaviour in wood poly (lactic acid) composites. Compos Part A Appl Sci Manuf 133:105871
15. Porebski A, Truong Hoang V, Vandenbroucke N, Hamad D (2020) Combination of LBP bin and histogram selections for color texture classification. J Imag 6(6):53
16. Raji A, Thaibaoui A, Petit E, Bunel P, Mimoun G (1998) A gray-level transformation-based method for image enhancement. Pattern Recogn Lett 19(13):1207–1212
17. Shivashankar S, Kagale MR (2018) Automatic wood classification using a novel color texture features. Int J Comput Appl 180:34–38
18. Zhang Y, Xu C, Li C, Yu H, Cao J (2015) Wood defect detection method with PCA feature fusion and compressed sensing. J For Res 26(3):745–751

Software Defect Prediction Survey Introducing Innovations with Multiple Techniques

M. Prashanthi, G. Sumalatha, K. Mamatha, and K. Lavanya

Abstract The software is applied in various areas, so that the quality of the software is very important. The software defect prediction (SDP) is used to solve the issues in the software and enhance the quality. Even if SDP is very helpful in testing, predicting the defective modules is not always easy. Different problems impede smooth performance and use of model defect prediction. The prediction of software defects was an interest of investigation, as early stage prediction of defects improves software quality with reduced cost and effective managing of software. Researchers from different fields help to propose different approaches that help effectively and efficiently. A number of approaches, frameworks, methods, and modeling were proposed using different data sets, metrics, and assessment strategies, in order to remove unnecessary and erroneous details from defect-prone modules. Defects in software systems are common and may cause software users various problems. During the development of different methods, the most probable defect location in large code bases was quickly predicted. Prediction of software faults is an important and beneficial way of improving software quality and reliability. The ability to predict which components in a large software system will contain the most faults in the next release contributes to better management projects, including an early estimation of possible release delays, and a "correcting guide for improving the software's quality. The identification of bugs/defects at the early stages of the software life cycle reduces the software development effort needed. A lot of research in software fault prediction using machine learning methods has been advanced. There are mainly two problems in the prediction of software defects, dimensional reduction, and imbalances of class.

Keywords Software defect prediction · SDP techniques · SDLC · SLDeep · Support vector machine (SVM) · COSTE

M. Prashanthi (✉) · G. Sumalatha · K. Mamatha · K. Lavanya
CMR Engineering College, Hyderabad 501401, India
e-mail: prashanthi.m@cmrec.ac.in

1 Introduction

Nowadays, for every task, the need for software is rapidly growing [1]. The software is applied in various areas, such as traffic signal command, biopharmaceutical engineering, and banking systems, due to the progression of the network society. Hence, the quality of the software products is very important, and the important five quality aspects are effectiveness, maintainability, availability, understandability, and reliability [2]. The advanced software systems are very complex, which cause various harmful negative impacts on the robustness and reliability of the software applications [3, 4]. Therefore, the software testing process in required for every software development project, which is a costly and critical process to examine the effectiveness of the resulting product. The major points considered to test the software is the total number of staff, the time required to complete the testing, and total amount need for testing [5]. The reliability and quality of the software mainly depend on the software testing, so that it is a necessary part of the whole software development procedure [6]. Anyhow, the major problems of software testing are resource-intensive activity, required processing time, and the budget need for testing [7]. There are two ways to perform software testing, such as the linear and cyclical approaches. The waterfall model is used in the linear approach, and agile, iterative, and incremental models are used in the cyclical approach [8]. The methods used to solve the issues in software testing and enhance the quality of software are called the software defect prediction (SDP) techniques [1].

The mistakes in the software development process are called software defects, which cause collapse, failure, faults, endanger human life safety, and property [2]. SDP is used to find the errors in the software before distributing it, which ensures the software quality [9]. Also, it is the major part of software development and a very expensive and complex part [10]. Moreover, it helps in the identification of the potential bugs and guides the resources for debugging [11]. So that it is helpful for enhancing the reliability and quality of the software that leads to robust and safer software artifacts [12]. The main aim of the software quality assurance (SQA) is for controlling the software development lifecycle (SDLC) and to make sure the present model meets the expectations. SQA contains various applications, like SFP, software testing, and code walkthroughs [13–15]. There are various approaches developed in recent years, which are used in the software development process. It helps the developers to assign the finite resources testing for the defective module [16–18].

Different machine learning algorithms for software defect prediction have been shown in Fig. 1. The SDP is one of the important phases of the software testing in SDLC. It is used to identify the defects in the software, and the resources are efficiently used without violating the constraints. There are various existing methods are developed for the prediction of software defects, in which some challenges are identified.
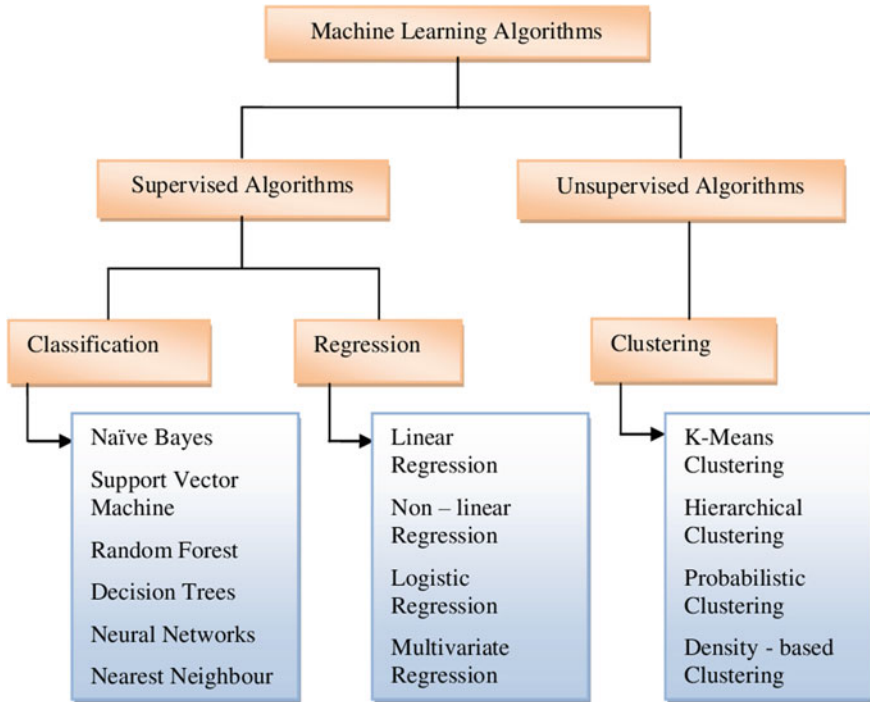
Fig. 1 Machine learning algorithms for SDP

2 Literature Review

The various existing methods used for the SDP are discussed in this section.

Majd et al. [1] developed a statement-level software defect prediction using deep-learning model (SLDeep) on static code features for predicting the software defects. This method minimized the hardships linked in pointing out the fault positions and also provided high-quality software with less effort and time. However, the effect of various nodes and layers, which are available in the long short-term memory (LSTM) network, was not exactly known.

Qiao et al. [3] devised a deep learning-based approach for predicting the defects in the software models. This method had high accuracy and attained maximum performance than the conventional methods. Anyhow, the total predictions from the change level were not known accurately.

Shi et al. [11] developed an unsupervised representation learning method for SDP, named multi-perspective tree embedding (MPT embedding). The program coding of this method was better than the conventional methods. However, this method was not used for some software, like code completion and clone detection.

Shao et al. [19] modeled a SDP model using the correlation weighted class association rule mining (CWCAR). This method was addressed the feature weights and

class imbalance of the SDP. Anyhow, this method was not applicable in real-world applications, like credit scoring and text classification.

Feng et al. [16] developed an oversampling technique named Complexity-based Over Sampling Technique (COSTE) to predict the defects in the software. This method was used to detect the diverse nature of synthetic instances, but it was not used in case of unsupervised and semi-supervised learning research.

Cai et al. [2] developed a hybrid multi-objective cuckoo search under-sampled software defect prediction model based on SVM (HMOCS-US-SVM). This method was used to remove the class imbalance (CIB) problem in data sets and used to select the parameter of SVM. However, some of the parameters of non-defective modules and SVM were not yet solved.

Ding and Xing [7] modeled a pruned histogram-based isolation forest method to predict the defects in the software. This method had a fast convergence rate and enhanced prediction performance using ensemble. The drawback of this method was the isolation features selection was carried out in a random way.

Sun et al. [20] devised a model to find the new defects data in the software, named collaborative filtering-based sampling methods recommendation algorithm (CFSR). This method was more effective and feasible in numerous cases to predict software defects. Anyhow, this method was not applicable in finding the differences between the homogeneous data, when the new data and historical data had similar software metrics.

Yucalar et al. [5] devised a Combining Predictors method, which is based on the multiple classifiers for the prediction of software defects. In this method, the performance was enhanced by reducing the effort needed for finding the software defects. However, this method was not compared with the empirical studies.

Zhao et al. [10] developed a cost-sensitive model, named Siamese parallel fully-connected neural networks (SPFCNN) for SDP. This method was used to solve the limited data and high dimensional problems, but it was not applicable for some forms of software defects.

Xu et al. [12] modeled Learning Deep Feature Representation (LDFR) for SDP. This method was useful in reducing the class imbalance problems of SDP. However, it was difficult to predict the reasons for defects in software modules.

Tumar et al. [8] devised an intelligent approach, named Binary Moth Flame Optimization (BMFO) with Adaptive synthetic sampling (ADASYN) for SDP. This method was useful in solving the class imbalance problems and finding the suitable features selection. Anyhow, the output was affected by the classifiers used in this method.

3 Different Models and Approaches

3.1 Soft Computing

Several soft computing methods have been recommended in the past for the prediction of software defects. Soft computing is a keyword of reference for aggregating different mechanisms related to computer science, such as AI methods, machine learning methods, and several other mechanisms which includes soft computing:

- "Artificial Neural Network"
- "Neural Network"
- "Support Vector Machine"
- "Swarm Intelligence: Ant Colony, Particle Swam Intelligence"
- "Machine learning techniques"
- "Probabilistic Reasoning"
- "Decision Tree"
- "K-Nearest Neighbors"
- "Evolutionary Computation"
- "Evolutionary Algorithm: Genetic Algorithm"
- "Fuzzy Logic"
- "Bayesian Belief Network".

Various statistical mechanisms such as the selection of feature subsets and PCA improve forecasting capacity in several SDP models. Software metrics play an important role, and models are supported in accurately predicting failures.

3.2 Software Metrics

Software metric is the software unit for measuring or specifying an attribute. These measurements are useful in determining software excellence. The quality metrics of software are a component of software metrics which focuses on product quality, methods and overall application features. Product metrics describe various characteristics of products such as volume, architectural design, computational completeness, efficiency, and quality. The process metrics of organizations are used to improve software development and support various tasks, such as product error detection, bug fixing during development, fault discovery during testing, and default removal time minimization. The project measurements define features and implementation of the projects involving software developers, recruitment patterns and price, project plan, and efficiency in the software's life cycle.

Various product metrics are

1. Chidamber and Kemerer.
2. The quality oriented extension to Chidamber and Kemerer metrics suite suggested by Tang et al.
3. Cohesion in Methods (LCOM3) suggested by Henderson-Sellers.
4. On the basis of McCabe's complexity metric the class level metrics built.
5. Martin suggested coupling metrics.
6. The Bansiya and Davis recommended QMOOD metrics suite.
7. Lines of Code (LOC).

3.3 A General Defect Prediction Process

In order to develop a prediction model, deficiency and measurement data from the software development efforts must be collected for use as a learning set. There is a compromise between how well a model matches its learning set and its performance in predicting the addition of information sets. We should therefore assess the performance of a model by comparing the predicted deficiency of the modules in a test against the actual deficiency. A general defect prediction process has been shown in Fig. 2.
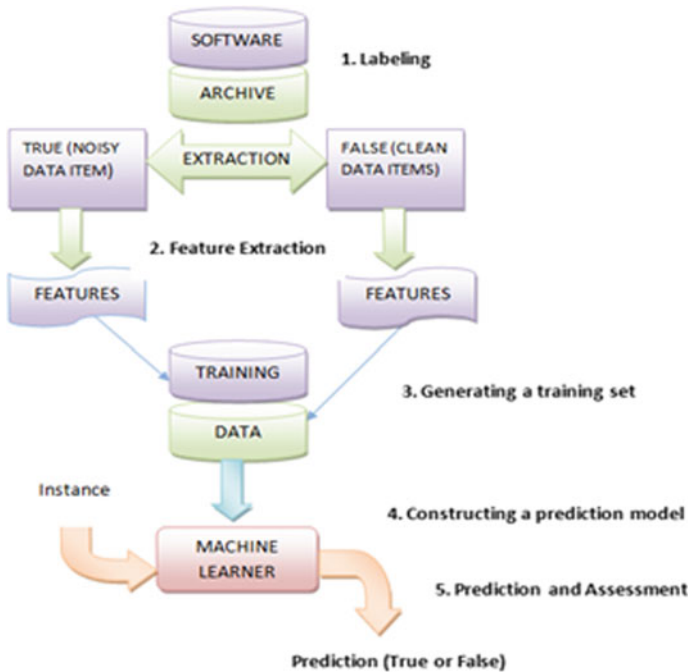
Fig. 2 General defect prediction process

3.4 Machine Learning-Based Models

Machine learning (ML) algorithms have proven to be very useful in solving a wide variety of technical problems including failure, error, and defect pulses, as the software of the system becomes more complex [21]. In cases where problem areas are not well defined, human knowledge is limited and dynamic adaptations are required to changing conditions for the development of efficient algorithms ML algorithms are very useful. Machine learning comprises various types of training, including artificial neural networks (ANNs), concept learning (CLs); Bayesian belief networks (BBNs); strengthened training (RL), genetic algorithms (GA) and genetic programming (GPs), and instance-based study (IBL) (AL).

3.5 The Fuzzy Logic Approach

The Fuzzy logic model builds on the concept or reasoning and develops an approximate value. It's a step forward from traditional Boolean logic, where True or False can exist only. The truth of any statement in the case of fuzzy logic is a degree rather than an absolute number. The greatest advantage of Fuzzy's logic, modeled upon the human intuition and behavior, is that this model does not answer traditionally, but rather gives an allotment to the human response.

The inputs of this model are placed in a series system. Then, a set of rules will be defining how inputs are used to get the output and to find the definitive value in the fuzzy set. The model has a set of metrics or RRML lists that are made of the available software metrics (Fig. 3). The measurements are relevant for their respective phases within the life cycle of software development.
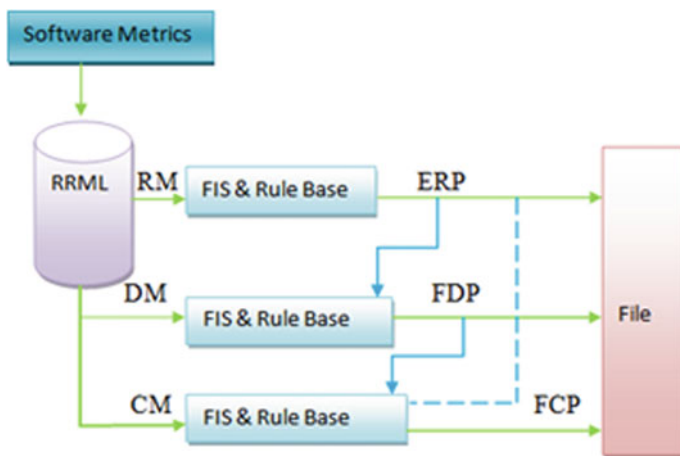
Fig. 3 Fuzzy logic approach

3.6 Code Pattern-Based Vulnerability Detection

There are two main steps in code pattern detection technology for vulnerabilities. During the workout stage, control and flow techniques are used to extract key program codes converted into a vector using the existing mainstream tool (e.g., word2vec) that can be used for supervisory training tools in suitable neural networks. During the detection phase, the same data processing is performed using a new software which identifies current vulnerabilities using the learned model. Code pattern-based vulnerability detection methods are divided into the current model training network structures of static detection methods and dynamic detection methods, including CNN, RNN, and LSTM.

3.7 Capture Recapture Analysis

This defect prediction technique is based on an assessment of patterns of defects found by independent defect detection activities in a particular software artifact. The count of latent defects is estimated by overlapping of defects identified by independent activities or groups of testers (the amount of defects remaining in a system). Default pooling is also called capture/recaptures techniques.

3.8 Expert Opinions

When experts are available, they use them for predictions based on their experience, the fastest and easiest method of defect prediction. The disadvantage of this methodology is its subjective nature and its inability to degrade correctly in lower granularity levels. This method may be useful if defect predictions at project or large component level are to be carried out, and if experts have experience of forecasting them, but if defect predictions are to be made at lower granularities (sub-systems, functions, files, etc.), this method does not scale down. This method should not be used.

3.9 Causal Model

Causal models try to establish causal relations with the expected number of defects or number of latent system failures, between software process and product attributes. Fenton and Neil criticize the application of statistical software defect prediction models for the lack of causal link modeling (BBNs). In very early stages in software projects, Bayesian Nets were used to demonstrate their application to defect forecasting.

3.10 Analogy-Based Predictions

Analog estimates are based on measurements between past and current projects collected and compared in an analogy in order to determine the most analogous project (s). Typically, size, application type, functionality complexity, and other parameters for predictions of a software defect are used to identify like projects for estimating. The analysis could be carried out at the level of the project, sub-system, or component.

3.11 Multivariate Regression

Models based on regression use statistical regression to make predictions of defects using software metrics or changing code as predictor variables. In a software project or modules (sub-systems/functions, etc.), multiple linear regression can be used to estimate the number of expected defects. As independent variables in regression-based models, a variety of software processes and product metrics have been used. Code complexity metrics and source code evolution (change) metrics are most common.

3.12 Constructive Quality Model (COQUALMO)

The model is used to build a quality model referred to as COQUALMO by expertly determined introductions and removals sub-models. The first estimated use of the sub-model Deficiency Introduction (DI) under this model was the number of non-trivial demands, design, and coding defects introduced. The DI sub-model uses the estimation of software size and other project and process attributes (platform, personal, etc.).

4 Automatic Software Program Repair

The primary purpose of early automating patching technology was to prevent worms from spreading that slowly penetrated all aspects of computer software safety through technological development.

The three phases of patching process are software failure, patch generation, and patch assessment. The location of software defects is a requirement for automatic repair and is primarily used to identify potential program defects or vulnerabilities. Common fault location technologies today are divided into two categories: static fault location technology and dynamic fault location technology. Data dependence and

data dependence relationship for program code tested is obtained by engineering for the location of the static defect generated mainly via program analysis technologies for confirming and locating the fault location. By running the default test case, a dynamic location technology receives information on the program execution and locates the default state position in the testing program by analyzing the execution flow of the program.

5   Conclusion

Early detection of software defects plays an important role in the software development cycle. In the automotive sector, development of software has largely adopted the model development paradigm that enables the easier integration of multi-provider functionality. Deficiencies are detected early and the intended functionality, robustness and compliance with model safety standards is verified and validated extensively—the quality and confidence of automotive software can be substantially improved. Effective approaches and instruments support cost reduction and reduction in development time. SDP is now dignified as a developing research zone using ML technologies. It is a challenging task to detect software failures during the first phase of SDLCS, as well as to finance high-quality software systems. The main highlight in this paper were several methods for predicting defects such as integrated approach, cross-project model, and machine learning algorithms. On the basis of the analysis, the best solution can be selected to analyze, predict, and avoid all mistakes and their limitations.

References

1. Majd A, Vahidi-Asl M, Khalilian A, Poorsarvi-Tehrani P, Haghighi H (2020) SLDeep: statement-level software defect prediction using deep-learning model on static code features. Exp Syst Appl 147:113156
2. Cai X, Niu Y, Geng S, Zhang J, Cui Z, Li J, Chen J (2020) An under-sampled software defect prediction method based on hybrid multi-objective cuckoo search. Concurr Comput Pract Exp 3(5):e5478
3. Qiao L, Li X, Umer Q, Guo P (2020) Deep learning based software defect prediction. Neurocomputing 385:100–110
4. Bowes D, Hall T, Petrić J (2018) Software defect prediction: do different classifiers find the same defects? Softw Qual J 26(2):525–552
5. Yucalar F, Ozcift A, Borandag E, Kilinc D (2020) Multiple-classifiers in software quality engineering: combining predictors to improve software fault prediction ability. Eng Sci Technol Int J 23(4):938–950
6. Lee SH, Lee SJ, Shin SM, Lee EC, Kang HG (2020) Exhaustive testing of safety-critical software for reactor protection system. Reliab Eng Syst Saf 193:106667
7. Ding Z, Xing L (2020) Improved software defect prediction using pruned histogram-based isolation forest. Reliab Eng Syst Saf 204:107170

8. Tumar I, Hassouneh Y, Turabieh H, Thaher T (2020) Enhanced binary moth flame optimization as a feature selection algorithm to predict software fault prediction. IEEE Access 8:8041–8055
9. Yu Q, Jiang S, Zhang Y (2017) A feature matching and transfer approach for cross-company defect prediction. J Syst Softw 132:366–378
10. Zhao L, Shang Z, Zhao L, Zhang T, Tang YY (2019) Software defect prediction via cost-sensitive Siamese parallel fully-connected neural networks. Neurocomputing 352:64–74
11. Shi K, Lu Y, Liu G, Wei Z, Chang J (2020) MPT-embedding: an unsupervised representation learning of code for software defect prediction. J Softw Evol Process e2330
12. Xu Z, Li S, Xu J, Liu J, Luo X, Zhang Y, Zhang T, Keung J, Tang Y (2019) LDFR: learning deep feature representation for software defect prediction. J Syst Softw 158:110402
13. Johnson AM Jr, Malek M (1988) Survey of software tools for evaluating reliability, availability, and serviceability. ACM Comput Surv (CSUR) 20(4):227–269
14. Wang H, Khoshgoftaar TM, Napolitano A (2010) A comparative study of ensemble feature selection techniques for software defect prediction. In: 2010 Ninth international conference on machine learning and applications. IEEE, pp 135–140
15. Hall T, Beecham S, Bowes D, Gray D, Counsell S (2011) A systematic literature review on fault prediction performance in software engineering. IEEE Trans Softw Eng 38(6):1276–1304
16. Feng S, Keung J, Yu X, Xiao Y, Bennin KE, Kabir MA, Zhang M (2020) COSTE: complexity-based oversampling technique to alleviate the class imbalance problem in software defect prediction. Inf Softw Technol 129:106432
17. Limsettho N, Bennin KE, Keung JW, Hata H, Matsumoto K (2018) Cross project defect prediction using class distribution estimation and oversampling. Inf Softw Technol 100:87–102
18. Nagappan N, Ball T, Zeller A (2006) Mining metrics to predict component failures. In: Proceedings of the 28th international conference on software engineering, pp 452–461
19. Shao Y, Liu B, Wang S, Li G (2020) Software defect prediction based on correlation weighted class association rule mining. Knowl Based Syst 105742
20. Sun Z, Zhang J, Sun H, Zhu X (2020) Collaborative filtering based recommendation of sampling methods for software defect prediction. Appl Soft Comput 90:106163
21. Song Q, Jia Z, Shepperd M, Ying S, Liu J (2010) A general software defect-proneness prediction framework. IEEE Trans Softw Eng 37(3):356–370

A Machine Learning and Fuzzy Heterogeneous Data Sources for Traffic Flow Prediction System

U. Mahender, Tattikota Madhu, and Rajkumar Patra

Abstract The information about Internet traffic should be accurate and timely important for various applications like admission control, congestion control, allocation of bandwidth, and anomaly detection. The prediction of traffic flow is vital for the management and policy of transportation. Mostly, earlier traffic flow prediction techniques utilized simple models for traffic prediction but still these techniques do not meet the desires of various applications of real world. To overcome this, machine learning and fuzzy heterogeneous data sources for Traffic Flow Prediction System (ML-TFPS) is designed and analyzed in this paper. Firstly, the time series model is utilized as a benchmark based on traffic data history for predicting the flow of traffic. Then, heterogeneous data will be integrated for Linear Regression (LR), extreme learning machine (ELM) with machine learning (ML) and fuzzy Traffic Flow Prediction System (MF-TFPS) model. To predict the features of traffic flow, Spark parallelization technology is utilized in described method. MF-TFPS will be intuitively visualized the results of traffic flow prediction. The MF-TFPS will be validated basing on the traffic flow of real data of San Francisco. Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) parameters will be utilized in this study for performance evaluation. From results it is clear that, MF-TFPS with RVM performs well in prediction of traffic flow than the LR, ELM models. The heterogeneous data will be more informative compared to the actual traffic data which is utilized by other researchers, and nonlinear technique utility is demonstrated that can resulting an improvement in the prediction accuracy of traffic flow.

U. Mahender (✉)
CMR Engineering College, Hyderabad, India
e-mail: u.mahender@cmrec.ac.in

T. Madhu
Nalla Narasimha Reddy Education Society's (NNRG) Group of Institutions, Hyderabad, Telangana, India

R. Patra
CMR Technical Campus, Hyderabad, Telangana, India

1 Introduction

For the last two decades, the urban traffic systems have been developed significantly
with the rapid and wild expansion of urbanization [1]. However, these significantly
increasing scale of traffic systems lead to many urgent issues like traffic accidents,
exhaust emission, and congestion. While considering these notorious issues adverse
impacts, the administrators of urban have achieved a consensus that vigilant and
accurate traffic flow prediction technique is essential for dealing the traffic issues
which are constructed mathematically and presented an Intelligent Transportation
System (ITS) [2].

The ITS is an integrated system that combined the advanced science and tech-
nologies like operational research, electronic information technology, control theory,
data communication technology, artificial intelligence, and sensor technology for
improving the industry of transportation [3]. The ITS researches and applications
involved water carriage, civil aviation, railway, highway, and other transportation
modes [4]. The prediction of traffic congestion is also called as prediction of traffic
flow state, and it is one of the essential parts of ITS. Unreasonable and untimely
prediction of traffic congestion leads to huge economical loss to society because
of increasing exhaust gas pollutions and reduction in living standards of citizens
[5]. Hence, prediction of traffic is of great value for traffic management. Generally,
prediction of Internet network traffic is crucial for computer network management
and for network providers.

The service providers of Internet might utilize the timely and accurate information
of traffic flow for allocating the bandwidth, detecting the anomaly, controlling the
congestion, and designing the networks. It has potential for helping the managers
of internet to make better scheduling decisions, improve the operation efficiency of
Internet and alleviation of traffic congestion. The aim of predicting Internet traffic
flow is providing the information of such Internet traffic flow. The prediction of
Internet traffic flow has received lot of attention while geographically distributed
and large scale applications are on rising [6].

The prediction of Internet traffic flow hugely relies on real-time traffic and histor-
ical data extracted from several monitoring sources of Internet flow. The traffic data
have been exploding with the wide spread classical traffic sensor and novel traffic
sensor technologies and entered into the Internet traffic and big data era. The Internet
traffic control and management driven by big data has become a new trend. However,
various Internet traffic flow prediction models and systems are also available; many
of them utilize shallow traffic systems, but still they are somewhat unsatisfied. This
inspired many researchers to rethink about the prediction of Internet traffic flow
based on deep learning methods with huge volume of Internet traffic data. The traffic
welfare contains certain factors that include parking costs, accident costs, passage

time, service costs, accessibility, and punctuality, while most significant feature is congestion of traffic [7].

As a part of ITS, the traffic surfaces are positioned in the place for collecting the information of traffic overall the roads in timely for providing the traffic information to users that includes congested areas, amount of traffic, and traffic accident locations. In this manner, the ITS improves the road traffic network functionalities. In addition, ITS provides a real-time information services. The congestion of road will decrease and traffic can be dispersed by an optimal path to every driver. Thus, an ITS mainly focused on immediacy but attains very less accuracy. For solving these issues, active researches have been performed for predicting the features of traffic congestion and determining the MF-TFPS that will evaluate short-term congestion of traffic in the future through the feature prediction results of traffic congestion.

The rest of this research work is organized as: Literature survey is discussed in Sect. 2. Framework of described model introduced in Sect. 3. The experiment results are described in Sect. 4. At last, conclusion and directions of future research are provided in Sect. 5.

2 Literature Survey

Habtemichael and Cetin [8] presented a data-centric, nonparametric method for achieving short-term traffic predictions depend on similar traffic patterns identification by improved K-nearest neighbor (K-NN) algorithm. The weighted Euclidean distance is also utilized as a measurement of similarity to K-NN. In this research, 24 datasets from the highways of US and 12 datasets are utilized from the highways of UK. 22% of MAPE rate is achieved.

Zhang et al. [9] extracted volume of traffic, density of traffic flow, and velocity of traffic flow for judging the level of traffic congestion through the grey relational membership degree rank clustering algorithm.

Lv et al. [10] presented a traffic flow predicting technique based on deep learning that inherently considers temporal and spatial correlations. Real time and accurate traffic information is vital for successfully deploying the ITS. The stacked auto-encoder technique is utilized for learning the features of generic traffic flow and trained in greedy layer wise. Further, the results have demonstrated that the presented model has superior performance in terms of traffic flow prediction.

Periyanayagi and Sumathy [11] presented a time series approach named S-ARMA (autoregressive-integrated-moving-average), utilizing ARMA and Swarm intelligence to predict the network traffic in wireless sensor networks. Various ARMA variants are presented for improving the accuracy of prediction like seasonal ARIMA (SARIMA), KARIMA (KohonenARIMA) and spacetime ARIMA. All kinds of time series models are utilized for the prediction of traffic flow except the ARIMA-like time series models.

Huang et al. [12] presented deep infrastructure that contains multitask regression layer and deep belief network (DBN) for forecasting the flow of traffic. This presented architecture contains two parts: multitask regression layer at the top and DBN at bottom. DBN algorithm can be utilized for supervised feature learning. The experiments on the datasets of transportation exhibit better performance. The experiment results demonstrated that this architecture has attained 5% improvements on state of art model. It can also represent that the MTL may improve the shared tasks generalized performance. The positive results described that MTL and deep learning is one of the promising approaches in transportation research.

Comert and Bezuglov [13] developed a technique to predict the parameters of traffic during abrupt changes based on change point modules. An intuitive technique employs Hidden Markov Model (HMM) and expectation maximization (EM) algorithms as a change point models at these shifts and appropriately adapt the ARIMA forecasting technique is formulated. It is compared with mean and basic updating forecast techniques. In detail, numerical experiments are provided on many days of data for representing the impact of change point models usage to adapt forecast models.

Chen [14] presented an ARIMA-GARCH (Generalized Autoregressive Conditional Heteroscedasticity) model for predicting the traffic flow. This approach combines the nonlinear GARCH model and linear ARIMA model, thus it captures the traffic flow series conditional heteroscedasticity and conditional mean. This hybrid model performance is compared with standard ARIMA approach. Results have demonstrated that the inclusion of conditional heteroscedasticity does not bring sufficient improvement for accuracy prediction in certain scenarios, the GARCH (1, 1) model might deteriorates the performance. Hence, standard ARIMA is suitable in ordinary traffic flow prediction.

Ghosh et al. presented a different time series technique classes named STM (structural time-series) approach (in its multivariate form) for developing a parsimonious and computationally simpler multivariate short-term traffic forecast algorithm. A study is conducted at Dublin, Ireland, city center with huge congestion of traffic for evaluating the forecasting mechanism. The results indicated that presented forecasting algorithm is an effective one for traffic flow prediction in real time at multiple junctions in the urban network transportation.

3   Traffic Flow Prediction System

The architecture of the machine learning and fuzzy heterogeneous data sources for Traffic Flow Prediction System (MF-TFPS) is represented in Fig. 1.

Based on traffic historical data, the time series models will hardly reflect the non-recurrent events; environmental and search engine data could be utilized as supplementary data sources. The reason behind this data collecting technique is simple and robust: Human transportation needs are driven through daily events.
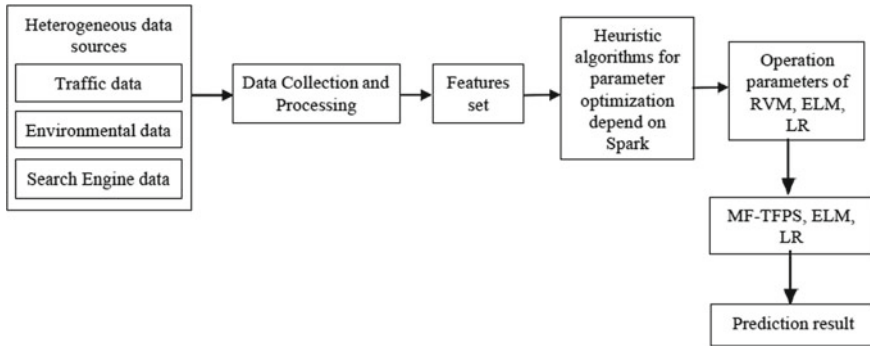
Fig. 1 Architecture of traffic flow prediction system

Since the Internet has been become an essential part in daily life, search engine data is treated as main index in daily events [5].

With these increasing societal concerns toward the environment, the level of air pollution will have a significant impact over the behavior of individuals. For satisfying the demands of prediction models, main important one is before the construction of model for preprocessing all the data. Mostly the real-world data is general, results incomplete attribute values, incorrect data or outliers, and inconsistent data formats. The preprocessing of data is a generic term that represents the transforming process of unstructured data into ease handling format for latter processing.

Multiple ways are available for conducting data preprocessing that involves data cleaning, transformation of data, integration of data, and data reduction. The next step is feature selection for improving the models overall performance and only requires less amount of data significantly. Feature selection can be used for identifying and finally redundant, and irrelevant attributes will be removed that have small or no contribution to the prediction model accuracy and even decrease its performance.

Traffic data with a matrix is applied as an input to single LSTM layer for extracting the features of traffic data. The optimizers with adaptive learning rate would perform better because of the traffic data sparsity.

When a neural network would be trained with small dataset, then it can be easy for causing overfitting. For preventing the overfitting, the neural network performance will be improved while adding a dropout layer in neural network. In LSTMSPRVM, 5% of neurons are disconnected in this dropout layer. The extracted features are collected at the output through a dense layer. Eventually, the set of new features are applied as input to RVM for prediction. If the traffic data sparsity is high, then the prediction difficulty is also high. However, the very short intervals can result in decision-makers do not have time for making accurate decisions in a less period of time.

The optimization of single kernel function parameters has poor capabilities if the samples are distributed unevenly in higher dimensional space. Hence, the functions of single kernel are combined linearly for the construction of combined kernel function

which is expressed as

$$f(x, x_i) = \exp\left(-\frac{\|x - x_i\|}{2\sigma^2}\right)\lambda + (1 - \lambda)\left[\gamma(xx_i + 1)^d + C\right] \qquad (1)$$

where λ defines the combined kernel function width and d indicates data distribution in higher dimensional space. λ is the coefficient of weight and meets $0 \leq \lambda \leq 1$.

At present, most generally utilized parallel computing frameworks are Apache spark and MapReduce. While in task execution, Spark utilizes multi-thread model whereas MapReduce utilizes multi-process model. However, spark depends on memory computing and mostly suitable for ML algorithms, stream computing, and iterative algorithms.

As the heuristic algorithms can be equivalent for iterative optimization algorithms, so for parallelization spark will be selected. Linear Regression (LR), extreme learning machine (ELM), and RVM are the three main classifiers used in this method. The LR classifier is utilized for predicting the actual values in accordance with the continuous input independent variables. The main aim of LR is fitting the better line corresponds to relationship between independent and dependent variables.

The Single Hidden-Layer Feed Forward Neural Network (SLFN) is one of the significant classification types in ML. Simulated and experimental results exhibited that ELM algorithm has better performance and is more suitable for SLFNs. In ELM, hidden biases and input weights are generated randomly rather tuned. Output weights (linking hidden layers to output layers) determination is sample as determining the least-square solution for given linear system.

A framework named machine learning and fuzzy Traffic Flow Prediction System (MF-TFPS) model is presented for traffic congestion prediction and visualization. The RVM is another classifier utilized in this approach. Parameter optimization algorithms are added to RVM this in turn doubles RVM training time. A RVM parallel training technique is designed for reducing RVM training time based on the parallelism of parameters. Designing a predictive technique mostly researcher always looking for maximizing the number of correct predictions among all predictions. After fitting these models, many indices are utilized for evaluating these models overall performance. Finally, overall performance is refined and improved while tuning each of these models.

4   Experimental Results

The demonstrated model predictive performance is discussed with three ML classifiers LR, ELM, and MF-TFPS with RVM. The MF-TFPS framework utilizes search engine and environmental data as supplementary data sources. Machine learning fuzzy comprehensive data sources Traffic Flow Prediction System (MF-TFPS) would

be validated based on San Francisco real traffic flow data. This design main advantage is increasing the data security in presented framework and potential attackers attack cost.

For robustness, relative ratios 70%, 80%, and 90% between testing subset and training subset are considered for evaluating the performance of prediction. RMSE and MAPE are chosen for evaluating the performance of prediction in terms of accuracy level. RMSE and MAPE are used for evaluating the accuracy of prediction.

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{x_i - x_i^{\sim}}{x_i} \right| \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |x_i - x_i^{\sim}|^2} \tag{3}$$

where x_i the real value of the ith observation, N is the total number of observations in testing dataset, and $x_i^{\sim}$ is the ith observation result of prediction model which is to be evaluated. Table 1 represents the MAPE values of three classifiers as LR, ELM, and MF-TFPS with RVM.

The MF-TFPS with RVM model performed better than ELM and LR. If training set ratio is same, then MAPE of MF-TFPS with RVM is clearly smaller in the case where independent variables would be identical. Graphical representation of the MAPE values is shown in Fig. 2.

Another evaluation is RMSE when performance level of prediction is discussed. Figure 3 represents the values of RME for training dataset different ratios and models.

Table 1 MAPE value of different models in different training set

Ratios of training set (%)	MAPE value		
	LR (%)	ELM (%)	MF-TFPS with RVM (%)
90	7.2	5.4	3.1
80	6.8	4.9	2.7
70	6.1	4.2	2.4

Fig. 2 Performance comparison of different methods by MAPE criterion
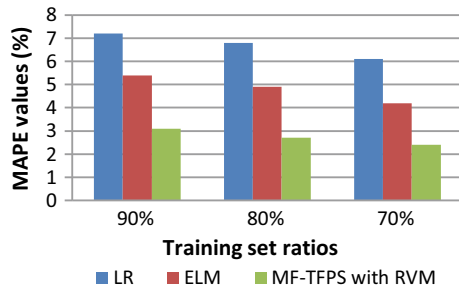
Fig. 3 Performance
comparison of different
methods by RMSE
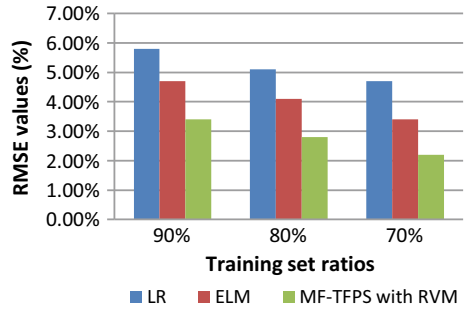
Table 2 RMSE value of different models in different training set

Ratios of training set (%)	RMSE value		
	LR (%)	ELM (%)	MF-TFPS with RVM (%)
90	5.8	4.7	3.4
80	5.1	4.1	2.8
70	4.7	3.4	2.2

It is clear that the obtained characteristics are similar like MAPE. Identical results
are observed. Table 2 represents the RMSE values of three classifiers as LR, ELM,
and MF-TFPS with RVM. Graphical representation of the RMSE values is shown in
below Fig. 3.

When the predicted value is near to true value, then the prediction approach
performs greatly (i.e., as the smaller RMSE and MAPE then performance is better).
Novel results are presented through modeling the environmental data, assembled
search volume index and real traffic flow for accurate traffic flow prediction. The
results demonstrated the external data advantages and exhibits that presented model
would achieve better prediction accuracy than the traffic data history collected by
sensors. As for model comparison, the MF-TFPS with RVM improves MAPE and
RMSE level. In different training environments, the result will slightly change that
demonstrated that while using the heterogeneous data MF-TFPS with RVM tends to
perform better overall.

5   Conclusion

Prediction of traffic flow is vital for the management and policy of transportation. In
this paper, architecture of machine learning and fuzzy heterogeneous data sources
for Traffic Flow Prediction System (MF-TFPS) is described. For RVM parameters
optimization, most of the heuristic algorithms have the concern that is speed of

convergence is very fast in iteration early periods this will lead to decrease of population diversity in the later period of iterations. Linear Regression (LR), extreme learning machine (ELM) with machine learning and fuzzy Traffic Flow Prediction System (MF-TFPS) with RVM model are used in this paper. MF-TFPS is validated based on the traffic flow of original data of San Francisco. MAPE and RMSE parameters are used in this study for performance evaluation. From results, it is clear that MF-TFPS with RVM performs well in prediction of traffic flow than the LR, ELM models.

References

1. Kamdar A, Shah J (2021) Smart traffic system using traffic ow models. In: 2021 international conference on artificial intelligence and smart systems (ICAIS)
2. Gao J, Gao X, Yang H (2020) Short-term traffic flow prediction based on time-space characteristics. In: 2020 IEEE 5th international conference on intelligent transportation engineering (ICITE)
3. Kotsi A, Mitsakis E, Psonis V (2020) Coordinated provision of C-ITS services for dynamic traffic management. In: 2020 IEEE 23rd international conference on intelligent transportation systems (ITSC)
4. Du Y, Man KL, Lim EG (2020) Image radar-based traffic surveillance system: an all-weather sensor as intelligent transportation infrastructure component. In: 2020 international SoC design conference (ISOCC)
5. Di X, Xiao Y, Zhu C, Deng Y, Zhao Q, Rao W (2019) Traffic congestion prediction by spatiotemporal propagation patterns. In: 2019 20th IEEE international conference on mobile data management (MDM)
6. Liu D (2019) Research on perceptual information fusion and traffic prediction based on internet of vehicles. In: 2019 international conference on robots and intelligent system (ICRIS)
7. Muthumanickam G, Balasubramanian G (2017) A traffic congestion control in urban areas with vehicle-infrastructure communications. In: 2017 international conference on energy, communication, data analytics and soft computing (ICECDS)
8. Habtemichael FG, Cetin M (2016) Short-term traffic flow rate forecasting based on identifying similar traffic patterns. Transp Res C Emerg Technol 66:61–78
9. Zhang YY et al (2016) A method for traffic congestion clustering judgment based on grey relational analysis. ISPRS Int J Geo-Inf 5(5):71–86. https://doi.org/10.3390/ijgi5050071
10. Lv Y, Duan Y, Kang W et al (2015) Traffic flow prediction with big data: a deep learning approach. IEEE Trans Intell Transp Syst 16(2):865–873
11. Periyanayagi S, Sumathy V (2014) S-ARMA model for network traffic prediction in wireless sensor networks. J Theor Appl Inf Technol 60:524–530
12. Huang W, Song G, Hong H, Xie K (2014) Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE Trans Intell Transp Syst 15(5):2191–2201
13. Comert G, Bezuglov A (2013) An online change-point-based model for traffic parameter prediction. IEEE Trans Intell Transp Syst 14(3):1360–1369
14. Chen C (2011) Short-time traffic flow prediction with ARIMAGARCH model. In: 22nd IEEE IV, Baden-Baden, Germany, pp 607–612

A Survey on Wireless Channel Access Protocols

Md. Gulzar, S. Kiran Kumar, Mohammed Azhar, and Sumera Jabeen

Abstract Networks are playing very important roles in human life for day-to-day communication tasks. The first network was wired. The medium in this type of network was cables. The drawback of these types of networks is huge, and cabling is needed which is cost-effective and it is difficult to move the device from one location to other. But present networks include mobile devices which need to be connected to the Internet wherever they are. For this type of devices, wired networks are not suitable. For mobile devices, we use wireless networks, for which the transmission medium is air and data is transmitted in the format of radio signals. Many devices can access the medium at the same time. So, we need protocols to make this wireless communication possible. These protocols are known as multiple access protocols, which are discussed in this paper.

Keywords Bandwidth · ALOHA · Channelization · Orthogonal · Non-orthogonal

1 Introduction

Wireless transmissions occur in shared medium using a shared medium and omni-directional antennas. The major drawback in wireless networks is the transmissions can collide with each other when there are so many pairs of transmitters and receivers communicating. If we make the communication between each pair as point-to-point link, we can improve the performance because there may not be collisions. But this will also include the problem of channel sharing. Wireless channel access protocols attempt to share channels among many pairs of transmitter and receiver. These protocols aim to increase the throughput of wireless channels by carrying many transmissions as well as controlling the collisions which may occur. To improve the performance and removing collisions, various multiple access protocols are proposed which are discussed in below sections.

Md. Gulzar (✉) · S. Kiran Kumar · M. Azhar · S. Jabeen
CMR Engineering College, Medchal, Hyderabad, India
e-mail: md.gulzar@cmrec.ac.in

2   Carrier Sense Multiple Access Protocol (CSMA)

This is a popular protocol. This implements "Listen before talk" mechanism. Each transmitter will sense the wireless channel before sending data. When channel is found to be idle, transmitter transmits the data. Collisions will occur if two or more transmitters found the channel idle and transmit the data. These collisions will not noticed by transmitters at the same time due to their positions are different from noise. Explicit ACKs or NACK are required to know about the collisions from intended receivers. CSMA is only collision detection protocol.

3   Carrier Sense Multiple Access Protocol with Collision Avoidance (CSMA-CA)

CSMA-CA has three parts

(a) Carrier Sense—Transmitter senses the channel before sending the data. It transmits only if channel founds to be idle.
(b) Multiple Access—Many transmitters and receivers can share the same channel.
(c) Collision Avoidance—More than one participant should not start a transmission at the same time to avoid collisions. If collision occurs, transmission will be stopped and restarts again.

CSMA-CA uses Distributed Coordination Function (DCF) to provide contention free channel.

DCF works as follows

(a) When the node senses the channel and founds it to be idle, it transmits the packet after determining the channel is continuously idle for duration of time known as Distributed Inter-frame Space (DIFS). The time duration required to process received signal and to generate response frame is known as Short Inter-Frame Space (SIFS). The SIFS value for IEEE 802.11a, 802.11n and 802.11ac (at 5 GHz) is 16.
(b) DIFS is calculated as follows

$$\text{DIFS} = \text{SIFS} + (2 * \text{Slot time})$$

(c) When the channel is sensed busy, transmitting node waits till the duration of DIFS for the channel is to be sensed idle. The node then chooses a random amount of time known as random back off time value from a back off window {0, …, CWmin-1}, where CWmin is known as the initial back off window size. Its value is decremented each subsequent idle time slot and freezes the decrementing count-down process when the channel is sensed busy, and reactivates the counter after the duration of DIFS if the channel is idle.

(d) The sender node transmits its packet when counter reaches zero, and then waits for an ACK from receiver. If the ACK is received within a duration of SIFS, the transmitting node identifies that its packet is successfully received by the destination node.

(e) If the ACK is not received within a SIFS duration, it is known as ACK timeout or if sender senses the transmission of another packet on the channel, it will restart the back off process at step c with increased random back off value chosen from an increased back off window, with double window size.

4 ALOHA

ALOHA is the earliest Random Access Method, which was designed for wireless LAN, but it can also be used in shared medium. It has two variations pure ALOHA and slotted ALOHA.

4.1 Pure ALOHA

Pure Aloha is a simple protocol, in which transmission medium is shared between all stations. When stations have frames to send, they will send the frames on shared channel. There is high possibility of collision when more than one station is transmitting at the same time as shown in Fig. 1.

From the figure, three stations are sharing a channel, there are two collisions. In first collision, all stations are involved, and in second collision, station 2 and 3 are involved.
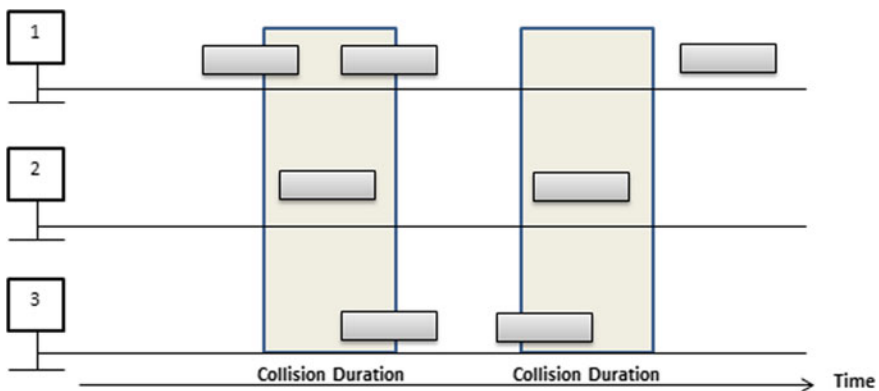
Fig. 1 Pure ALOHA

Pure ALOHA depends on acknowledgements (ACK), and sender considers the frame is successfully sent only after ACKs are received from receiver within a time-out period. When collisions occur, senders will not get ACKs within time-out period. If two or more stations timeout completed at the same time and if they retransmit at the same time again collision will occur. To overcome this, the each station will wait for random back off time after time-out period and then retransmits the frames. Like this, the collisions may reduce.

4.2 Slotted ALOHA

It is invented to reduce number of collisions in more effective way. In slotted ALOHA, the transmission time is divided into slot. Each station will send the data frame at the beginning of time slot. If a station has data to send and misses the time slot, it has to wait till the beginning of next time slot as shown in Fig. 2.

As shown in the figure, there are two collisions in slot 3 and slot 5, remaining transmissions in slot 1, slot 2, slot 4 and slot 6 will be succeeded.

5 Channelization

This technique is also called as channel partition where bandwidth of channel is shared in frequency, time and through code among stations. These protocols include FDMA, TDMA and CDMA.
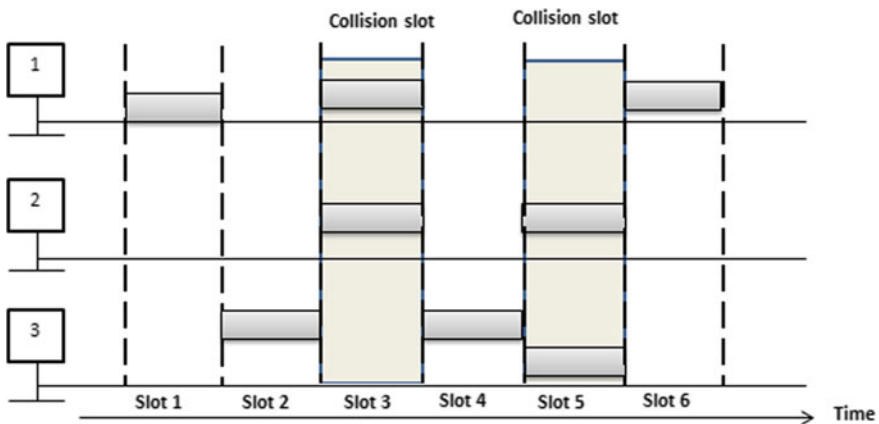
Fig. 2 Slotted ALOHA

Fig. 3 FDMA

5.1 Frequency Division Multiple Access (FDMA)

FDMA is a data link protocol. In FDMA, available bandwidth of transmission medium is divided into channels called as frequency bands. Each station will be assigned a band to send data through it for a long time. Each station uses a band pass filter to limit the frequency of transmission to that allocated channel. To prevent interference between adjacent channels, guard bands are used to separate the channels as shown in Fig. 3.

5.2 Time Division Multiple Access (TDMA)

In this technique, time is divided into slots. Each station is allocated one time slot. During the time slot, the station can use complete bandwidth of transmission medium. Each station must know beginning of time slot. This is difficult because of propagation delays of previous stations. To compensate the delays, guard time slots are used as shown in Fig. 4.

5.3 Code Division Multiple Access (CDMA)

As shown in Fig. 5, in CDMA, each station can access the transmission media completely and for complete time, i.e. there is no time sharing. In CDMA, each station has a code and data, for example, for Station 1, the code is c1 and data is d1. All the stations that are willing to transfer code on transmission media pass the product of code and data on transmission media. For example, S1 passes c1.d1, S2 passes c2.d2, S3 passes c3.d3 and S4 passes c4.d4. The transmission media takes the sum of all these products. That is transmission media has c1.d1 + c2.d2 + c3.d3 + c4.d4. If any station needs to receive data from any other station, it multiplies the
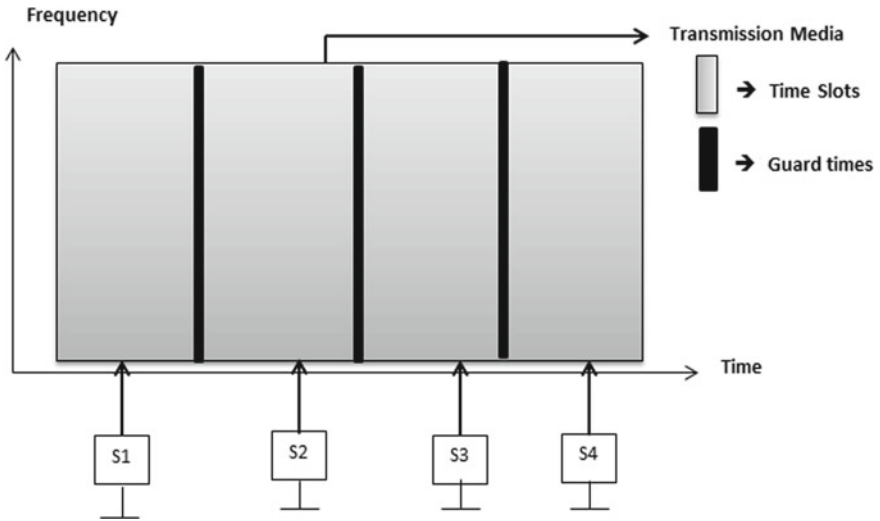
Fig. 4 TDMA

sum on channel with code of sender and divides the complete result by 4. Here, two properties are followed.

1. If each code is multiplied by other code, the result is 0.
2. If code is multiplied by same code, the result is 4.
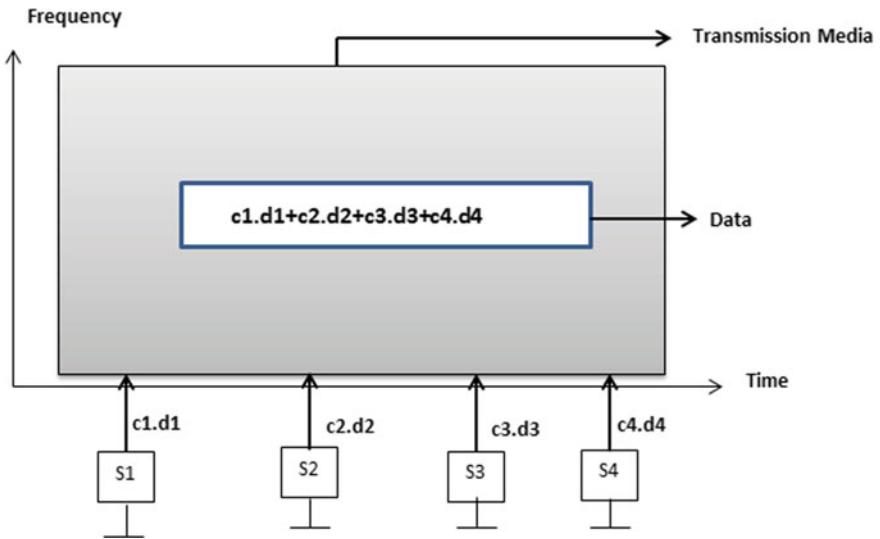
Fig. 5 CDMA

According to these properties, if station 3 what to receive from station 1, it multiplies sum of products on transmission media with code of station 1, i.e. c1 and divides the result by 4 as follows

- ((c1.d1 + c2.d2 + c3.d3 + c4.d4).c1)/4
- (c1.d1.c1 + c2.d2.c1 + c3.d3.c1 + c4.d4.c1)/4(4d1)/4
- d1.

6 Orthogonal Frequency Domain Multiple Access (OFDMA)

It is like other multiple access protocol which base station communicates with many mobile devices at the same time. OFDMA is used in LTE, IEEE 802.11a, 802.11g, 802.11n and in WiMAX (IEEE 802.16).

In OFDMA, the transmitter converts the outgoing stream in to symbols and sends the symbols on different frequency channels known as subcarrier. The bandwidth of subcarrier is small but collectively the bandwidth is equal to traditional single carrier system (Fig. 6). The space between two sub carriers is denoted subcarrier spacing Δf. The relationship between Δf and duration of each symbol T is given as follows

$$\Delta f = 1/T$$
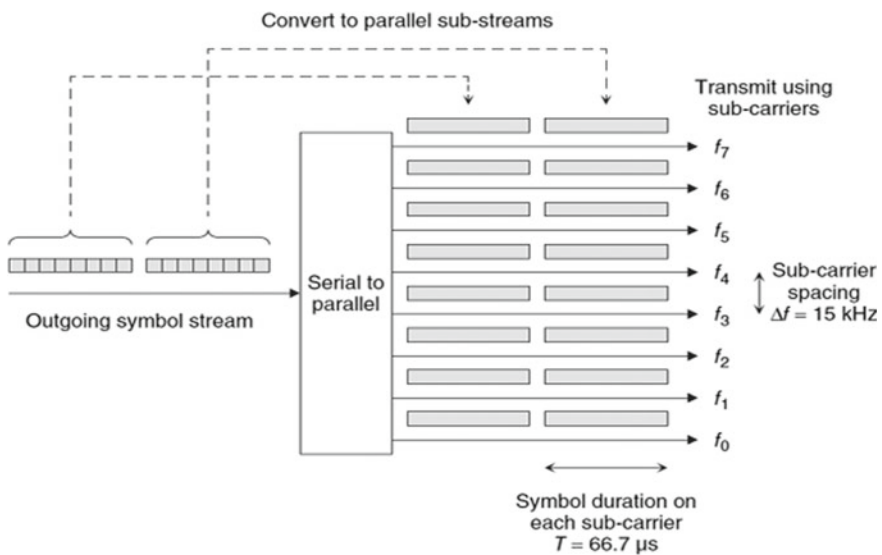
Fig. 6 Subcarriers in frequency band

where symbol duration in LTE is 66.7 µs and Δf is 15 kHz. Each cell will have 1200 subcarriers, which will occupy the central 18 MHz bandwidth of a 20 MHz allocation.

7   Spatial Division Multiple Access (SDMA)

SDMA is an advanced technique which increases bandwidth usage and spectral efficiency. Traditional systems will radiate power in all directions using omni-directional antennas. This causes interference in adjacent cells and wastage of power. The incoming signal becomes weaker due to noise and interference [1]. Omni-directional spreading of signal can also cause security threats to the information. In SDMA, smart antenna technology is used to track the location of mobile devices. The base station can adjust the phase of signals from multiple antennas and steer a beam to or from each user. SDMA technique used in wireless broadband systems and can also be used in mobile WiMAX, LLTE, GSM and 5G. SDMA increases capacity of the system and quality of transmission by focusing the signal into narrow transmission beams and pointing the beam at the direction of mobile station using smart antennas. SDMA can serve many users in same region. Mobile stations outside the region of directed beams will have no interference from other mobile devices operating under same base station. SDMA narrow beams have larger coverage with less radiated energy [2].

8   Frequency Hopping Spread Spectrum (FHSS)

In this technique, a spread spectrum is generated for hopping (changing) carrier frequency. Hopping happens on frequency bandwidth which will consist of number of channels. The channel which is used for hopping is called as instantaneous bandwidth and hopping spectrum is called as total hopping bandwidth. FHSS uses a signal less than 1 MHz of narrow band. The data signal is modulated on this narrow band carrier signal which hops from frequency to frequency. Frequency hopping is changing of transmission frequency. Frequency hopping is of two types, i.e. slow hopping and fast hopping. In slow hopping, on same channel is used for more than one data symbol to be transferred, and in fast hopping, for one symbol, frequency changes many times. Hopping sequence means which next channel to hop. Interference at any channel will affect the signal for small duration only because the signal will change the channel by hopping [3].

9 Direct-Sequence Spread Spectrum (DSSS)

In this DSSS, data signal of sending node is combined with extra bits of data, which is known as chips. The chip bits are added to the data bits in a ratio known as spreading ratio. A more chip bits are added the data signal will become immune to interference. These chip bits which are redundant are also used to recover original data bits if transmission is corrupted from remaining chipping code. In DSSS, transmitter and receiver contain similar pseudo-random sequence generators for producing Pseudo-Noise (PN) signal. At sender side, before transmission input, data bits are XORed with PN signal and this data is transferred to the receiver. At receiver side, the received data will be XORed with receiver's PN signal to get the original data bits sent by transmitter. This data with redundant bits is transmitted on several frequency bands known as channels within a particular frequency band is in FHSS [4].

10 Power Division Multiple Access

Power Division Multiple Access (PDMA) is a MAC protocol which contains power division. In PDMA, the transmitted power is divided into power segments (PS) like time slots and subchannels in TDMA and FDMA. Multiple power segments are used concurrently to deliver information in same channel. PDMA works on a principle called as super position coding (SPC) [5]. The main idea of SPC is to combine many messages into a single signal of many layers. SPC uses degraded channel for communications like one-to-many and many-to-one, and in peer-to-peer communication, SPC uses ergodic channel. The PSs in PDM are needed to be equal for related information streams if information streams have same priority we require the bit-fairness. Hence, PDMA works on two steps [6].

11 Orthogonal Power Division Multiple Access (OPDMA)

This technique is based on PDMA. It will utilize degraded channel and multiplexing to gain energy savings. OPDMA uses orthogonal power segments (PS) [7]. The data can be delivered under various channel gains. OPDMA used in different channels along with QoS guaranty for multiple information streams. OPDMA protocol is used in one-to-many and many-to-one type of communications.

12 Non-orthogonal Multiple Access (NOMA)

NOMA is a radio access method used in cellular communications. It provides greater spectrum efficiency. NOMA includes different types like Power Domain and code domain. Super Position coding (SC) is used in Power Domain NOMA at transmitter, and at receiver, successive interference cancellation (SIC) is used [8]. NOMA can be used to fulfil the requirements of fifth generation (5G) technologies. NOMA is opposite of Orthogonal Multiple Access (OMA). In NOMA, many users at the same time within same cell will be allocated one frequency channel. The advantages of NOMA are improved spectral efficiency, higher cell-edge throughput.

13 Conclusion

Various channel access protocols are studied in this paper. Since every network application has different bandwidth demands and needs to use different channel access methods, we require the channel access methods which provides good spectral efficiency and provides low latency and which are secure form unauthorized access and which are fault tolerant as well as which provides good congestion and collision control mechanisms. Further in future, as the demand for wireless networks is increased, we may see new wireless channel access protocols.

References

1. Feng Z, Zhang Z (1998) Smart antenna and spatial division multiple access. IEEE. ISBN: 0-7803-4308-5/98
2. Roy RH (1997) Spatial division multiple access technology and its application to wireless communication systems. IEEE. ISBN: 0-7803-3659-3/97
3. Macleod MD (1993) Telecommunications engineer's reference book
4. Zhang P (2010) Advanced industrial control technology
5. Han W, Ma X (2016) Power division multiplexing. National Natural Science Foundation of China under Grants 61401320, 61501285, and 61401354
6. Mazzini G (1998) Power division multiple access. IEEE. ISBN: 0-7803-5106-1/98
7. Han W, Zhang Y, Wang X, Li J, Sheng M, Ma X. Orthogonal power division multiple access: a green communication perspective. IEEE J Sel Areas Commun. http://doi.org/10.1109/JSAC.2016.2600139
8. Riazul Islam SM, Avazov N, Dobre OA, Kwak K-S (2016) Power-domain non- orthogonal multiple access (NOMA) in 5G systems: potentials and challenges

Sahaay—A Web Interface to Improve Societal Impact

Kayal Padmanandam, K. N. S. Ramya, Ushasree Tella, and N. Harshitha

Abstract In our society, there are numerous people who are in need of essential day to day survival needs. We also have providers who are willing to help. But the providers could not find the right collaboration with NGOs and other non-profit organizations to help the needy. This is due to a clear absence of efficient interface between the people in need and people willing to contribute. Our application intends to focus on providing a medium for the users to contribute to the NGOs, senior citizens and physically disabled in small scales towards the bigger cause. This application is a two-user interface for the NGOs and the user. The two-end users of this application are the people in society and the organizations. It is a user-friendly website created using AngularJs and Django. MySQL database was used for storing data and performing manipulation. The website provides a platform for posting needs and for posting information regarding events. It provides an option for volunteers to contribute monetary and other help as requested. It allows volunteers to check their past contributions and an invoice will be provided for their contributions.

Keywords Web interface · NGO · AngularJs · Django · MySQL · Sahaay

1 Introduction

In a country, there are lakhs of people in need of help and some people are willing to contribute at least on small scale but are unable to collaborate with NGOs and other non-profit organizations due to a lack of medium to offer help. There is a clear absence of an interface between the people in need and people willing to contribute. Even if we try to contribute to NGOs online, there is always the problem of authenticity. There is also the problem of finding a cause that speaks to us and makes us contribute. So, this work is our attempt to bridge the gap between NGOs, volunteers and people in need of help. It helps in different ways instead of only monetary which sets it apart from crowdfunding platforms. It assigns the volunteers and tracks the progress of the

K. Padmanandam (✉) · K. N. S. Ramya · U. Tella · N. Harshitha
BVRIT HYDERABAD College of Engineering for Women, Hyderabad, India
e-mail: kayalpaddu@gmail.com

posted need to help meeting all the needs of people and all the posted needs are being worked towards fulfilling. This platform can also be used to post about various non-profit events to raise awareness or funds, etc.; hence, the volunteer or any user for that matter can RSVP on the website. It can also help in tracking individual progress since it can be used to track all the contributions by a user and invoices will be generated for their contributions. This enables them to encourage others to help and can be used for legal purposes as well. Today our society has many NGOs that are striving to improvise the society and help the people in need once in every possible way. But, the activities or events carried out by NGOs have very little exposure to the public as they tend to reach out to the people at a physical level. If these NGOs are provided with an online platform to create awareness for the events being held nearby or anywhere, it will enormously positively affect the participation of people helping to create a visible impact on society. One of its basic objectives is to bridge the communication gap between the NGOs and society. Moreover, the collaboration of multiple NGOs to support collective events and also sharing resources among them is yet another important feature of this online platform. Also, aspects such as transparency about NGO's activities will be maintained and available for users at any point in time, thus encouraging more and more people to be a helping hand and be a part of good positive change in society. The work focuses on bridging the gap between NGOs and volunteers connecting with people in need. It helps in communicating a need and help in solving the problem for people in need. A volunteer will be assigned to each need and the progress will be tracked to avoid starvation of a need. It also aims towards raising awareness about the importance of various crucial issues which require attention from society. It can also act as a platform for posting about events for raising awareness and funds and many other non-profit events. It can help people contribute what they can afford to donate instead of only monetary donations hence increasing the recycling culture among the users. Each user's contributions can also be tracked on their profile, and the user's contributions invoice can also be used for legal purposes. All these culminate towards the main aim of helping those in need and helping volunteers donate towards what they want, however they want.

2 Literature Survey

Mridula Goel, Aryan Agarwal, Namit Chandwani, Tanmay Dixit—Building an application framework to connect NGOs, and volunteers [1] aims to set up platforms to allow NGOs to communicate these volunteering opportunities and for volunteers to enrol for those that match their preferences. The platform is developed both as an Android application and a web-based interface. This paper also explores how gamification can be used to make the process fun and competitive to increase engagement with the volunteers.

Kunal Pardeshi, Asmita Bhogade, Rohit Kulkarni, Tejal Shigwan, Prof. Karan Mashal, a cross-platform application to enhance NGO and society collaboration [2] provides a common platform for NGOs and society to make use of connectivity for

social services, information management and also promote major social events of the government. This platform majorly serves as a common base for collaboration among organizations and every other common man. This system will also ensure safe and trusted use of donations by organizations and maintain transparency in its operations.

Titarmare, Neha and Krupal, Prajakta and Tol, Mahesh and Gupta, Aradhya and Kolte, Shreyas in [3] discussed the features of the proposed application include a list of non-profit organizations; this application helps by providing a list of non-profit organizations in the World. To simplify the donation, process a call, email, push notification feature is used in the application that can help build the non-profit organization system so the donor can provide more information on the donation and ask the non-profit organizations questions. Next, this application implements the layered architecture in the development process.

Janhavi Desale, Kunal Gautam, Saish Khandare, Vedant Parikh, Dhanashree Toradmalle in [4] aim to aid such NGOs and provide them with the necessary IT infrastructure to optimize the use of resources and increase their reach for food and money donations. The work includes a cross-platform mobile application that will help them manage their volunteers, get orders by spreading awareness through a social media module to connect to people who wish to support them in this noble cause. The literature survey of existing apps gave insights about designing the necessary modules that we must have.

Jose Ramon Saura, Pedro Palos-Sanchez and Felix Velicia-Martin in [5] discussed the aim of this paper is to use an extended Technology Acceptance Model (TAM) to analyse the acceptance of a technological platform that provides a point of contact for non-profit-making organizations and potential volunteers. The TAM is used to find the impact that this new recruitment tool for volunteers can have on an ever-evolving industry. The TAM has been extended with the image and reputation and visual identity variables to measure the influence of these non-profit-making organizations on the establishment and implementation of a social network recruitment platform. The data analysed are from a sample of potential volunteers from non-profit-making organizations in Spain.

3 System Analysis and Design

3.1 Proposed System

Sahaay web application is implemented using the MVT framework. The Model View Template (MVT) is a software design pattern as shown in Fig. 1. It is a collection of three important components: model view and template. The model helps to handle databases. It is a data access layer that handles the data. The template is a presentation layer that handles the user interface part completely. The view is used to execute the business logic and interact with a model to carry data and render a template. For
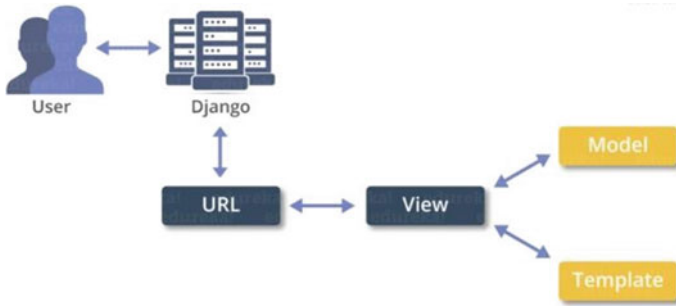
Fig. 1 MVT architecture

Sahaay, the front end was created using AngularJS, while the backend was Django framework along with the MySQL database. A template is a form of HTML that tells Angular how to render the component.

3.2 Architecture Design

This work has a two-end interface system with one end as a user and the other as a volunteer. Each of these has its own set of capabilities and actions. The figure above shows these capabilities and actions for each of the users. As depicted in Fig. 2, both end users can log in/register into the system. These needs can be accessed and seen by any user or volunteer at any given time as long as they open. These needs are displayed in the FIFO format on the home page. Users register for an event and attend the event. The user will receive a certificate for participating in the events. On the other end, NGOs can post information about the events and allow registered users to register. Volunteers can post about the events and also view the details about the participants who registered for the event. Volunteers can also take up needs posted by the users and connect with the user for more information about the need, etc. and get an invoice for the need they help fulfilled.

4 Implementation and Results

To implement the application, the following modules are needed. To implement those modules, front end, back end and database are required. The following are the modules, front end, back end and the database used.
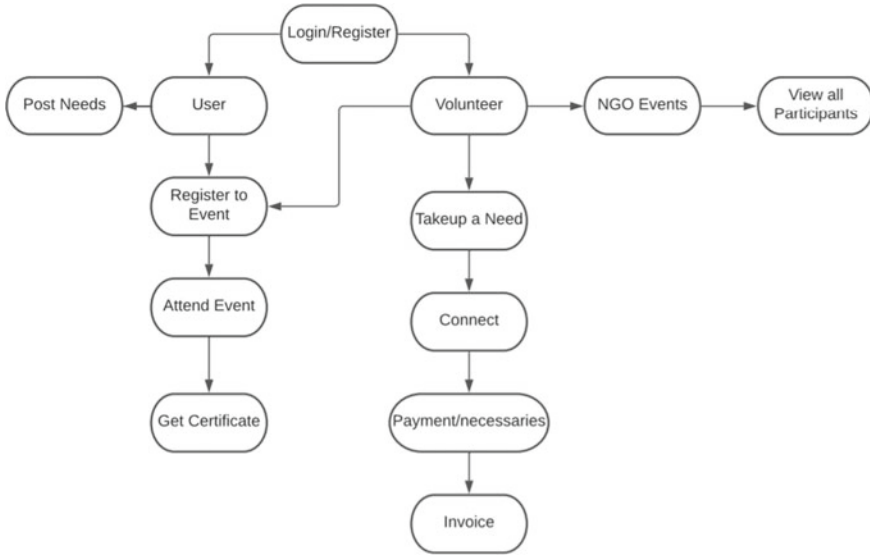
Fig. 2 Architecture design

4.1 Modules

Sahaay has the following modules

1. Volunteer module
2. User module
3. Needs module
4. Events module
5. Profile module.

4.2 Output Screenshots

The page contains the register button along with the functionality of the registration button, along with the options to navigate through the website using the different buttons on the page. The application pages are shown in Figs. 3, 4, 5, 6, 7, 8, 9 and 10.

5 Conclusion and Future Enhancements

Sahaay web application will help bridge the gap between NGOs and people in need. This will help people in need post their needs and others. It also supports the NGOs
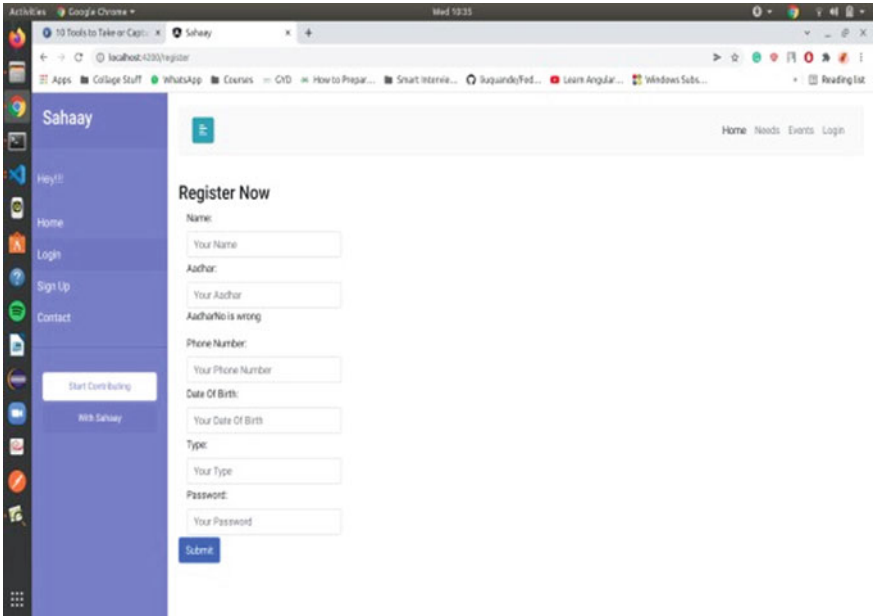
Fig. 3 Register page
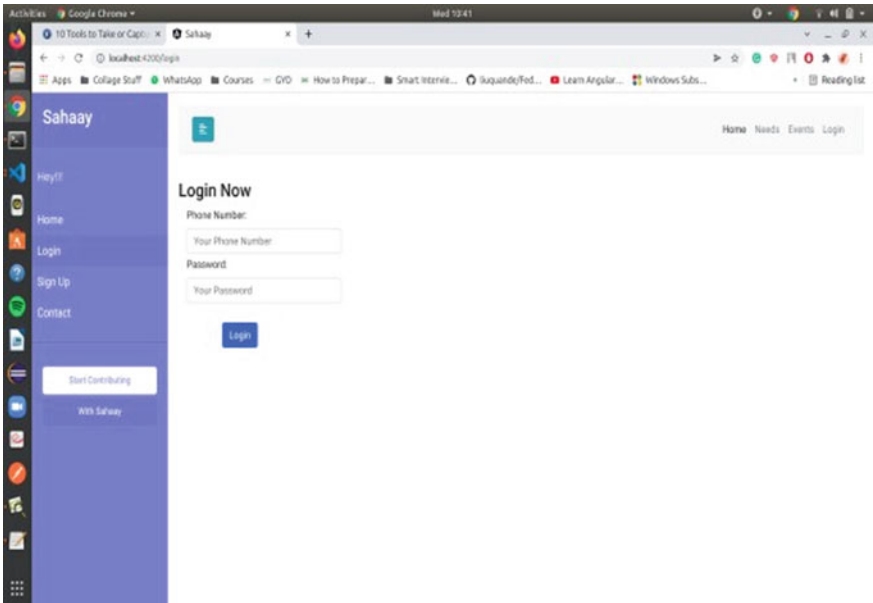
Fig. 4 Login page
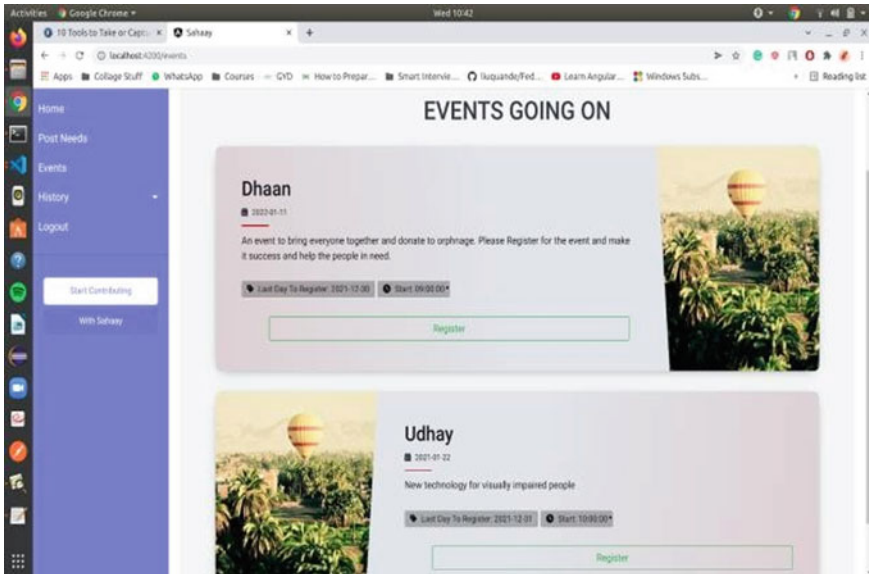
Fig. 5 Event page
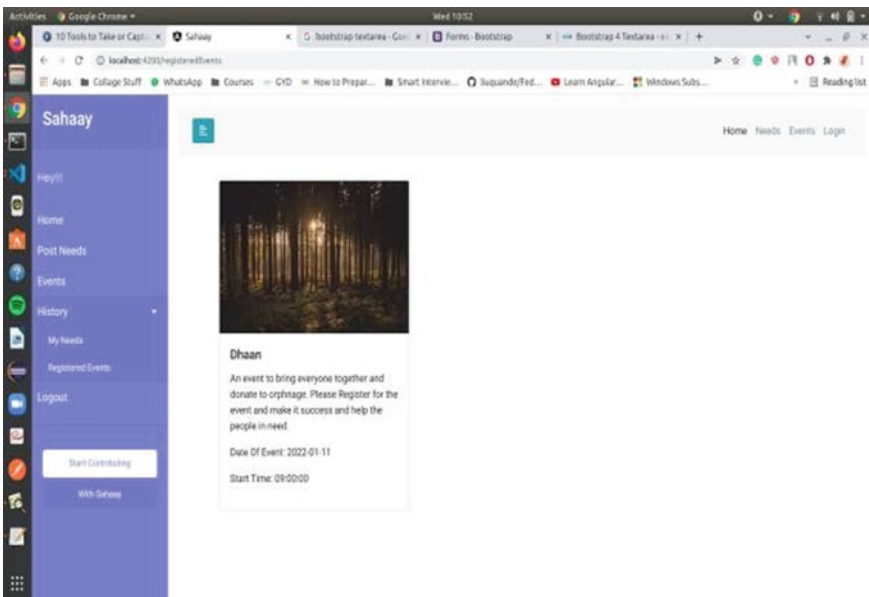
Fig. 6 Registered events page
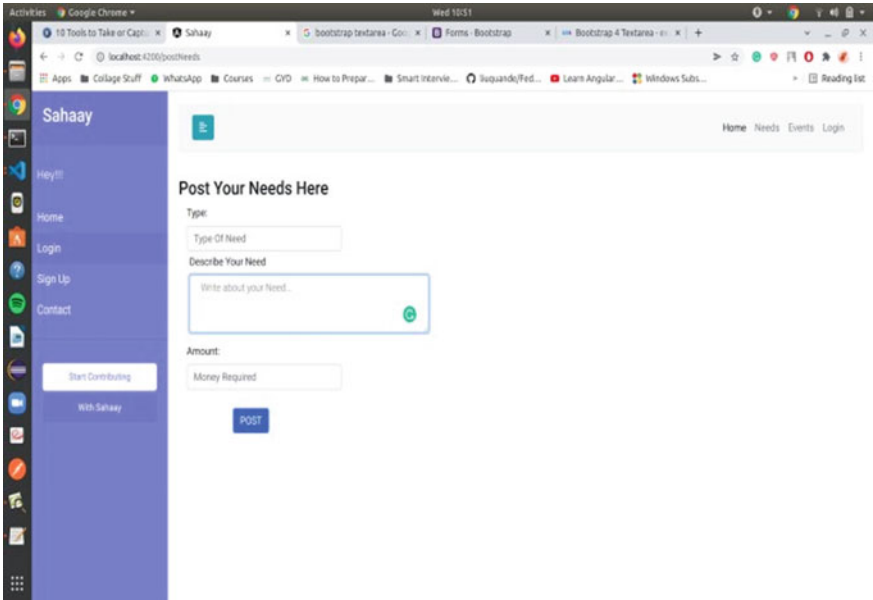
Fig. 7 Poster need
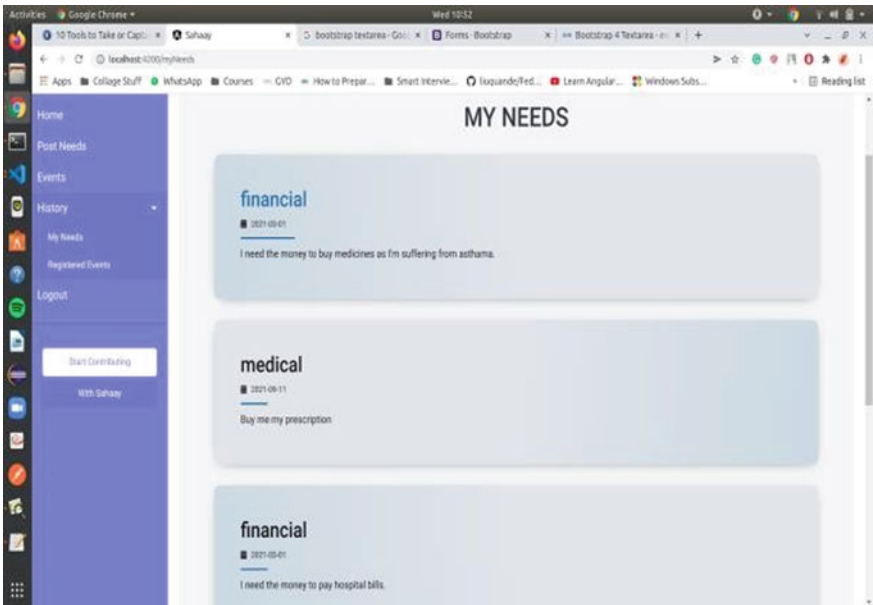
Fig. 8 My needs page
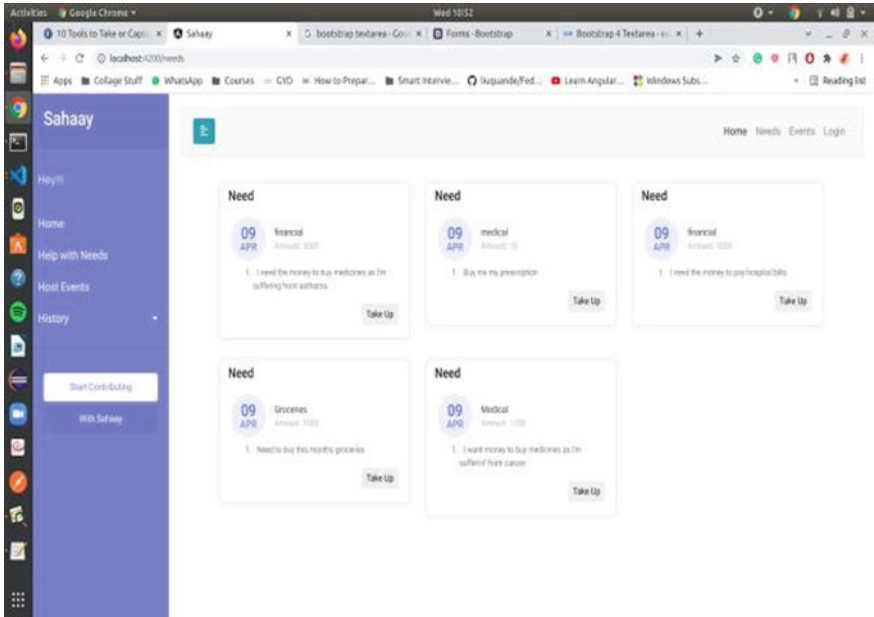
Fig. 9 Needs page
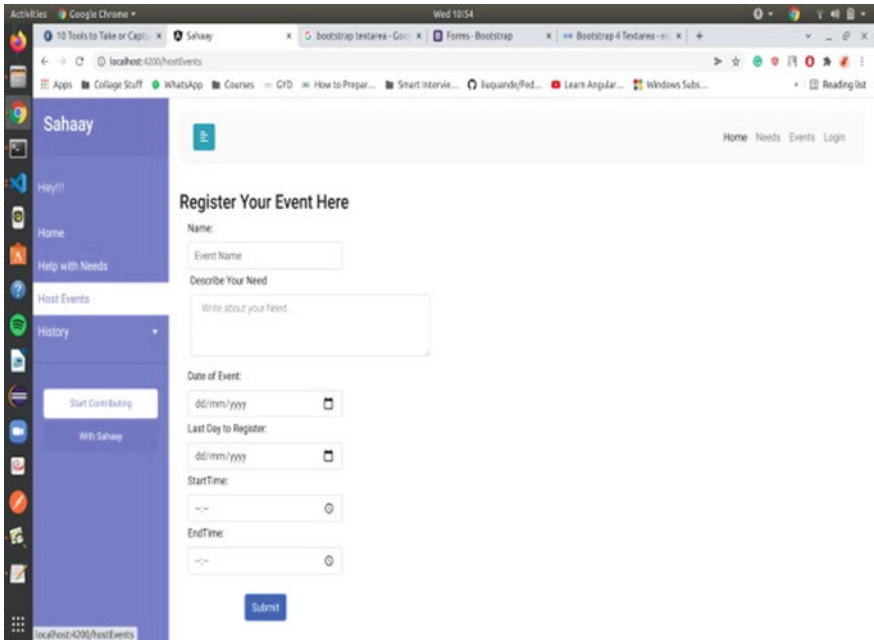
Fig. 10 Register events

by allowing them to post regarding the events organized by them. The user can post needs on this platform which can be fulfilled by volunteers. The needs and events which are posted on the platform are tracked by keeping the status updated hence avoiding starvations. All the users and NGOs are authenticated; hence, the user can be at ease while browsing through the application.

The web application can be up for further advancements by adding in-app payment features, needs to be provided with geolocation, chat window, etc. The posted needs and events can be displayed more interactively hence reaching a wider audience. The application aims to help bridge the gap between NGOs interacting and people in need, and this can be achieved by this application.

References

1. Goel M, Agarwal A, Chandwani N, Dixit T (2021) Building an application framework to connect NGOs and volunteers. In: International conference on innovative practices in technology and management (ICIPTM)
2. Pardeshi K, Bhogade A, Kulkarni R, Shigwan T, Mashal K (2016) A cross-platform application to enhance NGO and society collaboration. Int J Eng Sci Res Technol 5(4):75–78
3. Titarmare N, Krupal P, Tol M, Gupta A, Kolte S (March 7, 2020) Happy to Help (Hth): an android application and website for helping people to make donations. Int J Res Anal Rev (IJRAR) 7(1):1–6. E-ISSN 2348–1269, P-ISSN 2349–5138. Available at SSRN: https://ssrn.com/abstract=3677091
4. Desale J, Gautama K, Khandare S, Parikh V, Toradmalle D (2020) NGO support software solution: for effective reachability. Int J Educ Manage Eng (IJEME) 10(6):17–26. https://doi.org/10.5815/ijeme.2020.06.03
5. Saura JR, Palos-Sanchez P, Velicia-Martin F (2020) What drives volunteers to accept a digital platform that supports NGO works? Front Psychol

Summarization of Unstructured Medical Data for Accurate Medical Prognosis—A Learning Approach

Amita Mishra and Sunita Soni

Abstract The problem of "medical overload" has prompted research that includes a summary of automated medical documents and defines a system that will summarize medical transcript—raw and informal medical data. Medical transcript summarization has the power to save time, make notes, facilitate clinical decisions, and reduce medical errors by providing accurate treatment and medication. Physicians must combine information from different data sources to communicate with colleagues and provide integrated care. Medical documents contain a lot of data and are difficult to use in the analysis due to their unstructured format. Sequential structured data format for important details such as symbols and history. Medical details are needed to create better summaries and efficiency. Informal, irregular, or negative connotations can lead to failure to communicate and even medical errors. In this work, we will summarize medical transcript text, which aims to automatically predict the patient's required diagnosis based on clinical notes. In this approach, we use random data sampling data using an algorithm for abbreviating text to summarize diagnosis, prognosis, and treatment based on disease prediction and treatment by machine learning. The program will be based on more than 95% of diagnoses and more than 95% of symptoms and findings associated with each patient's diagnosis.

Keywords Medical transcript · Natural language processing

A. Mishra (✉)
CMR Engineering College, Hyderabad, India
e-mail: amita19092010@gmail.com

S. Soni
BIT-Durg, Durg, Chhattisgarh, India
e-mail: sunita.soni@bitdurg.ac.in

1 Introduction

When a patient is admitted, physicians should always spend more time reading the details of a medical record that may or may not be appropriate in the current situation. Medical professionals do not have enough time to read all the details but have to make critical decisions in time so they will benefit from an accurate summary of medical records. The medical record refers to the process of transferring a patient's voice record by a physician or health care staff during a patient consultation these documents are a physician or medical decisions and future reference, so it is very important to do this with high accuracy this software greatly reduces clinical document time and allows physicians that they are more focused on patient treatment programs. Remote access to the calling platform allows doctors to decide anywhere and anytime, leading to higher acceptance of medical records. Automatically converts medical and pharmacological guidelines used in doctor-prescribed notes, doctor–patient consultations, and telemedicine from speech to text for use in clinical document applications.

2 Literature Review

Xiangming Wang et al. proposed the machine learning techniques applied in the field of intelligent health analytics. The mechanism proposed in this paper is a random forest algorithm that targets continuous data and builds a predictive model based on regression analysis techniques. It can be applied to the analysis and prediction of control data in the intelligent medical industry [1].

Davide Chicco et al. proposed the system consists of a classifier that can read the electronic medical records of 18 of 22 patients diagnosed with sepsis and calculate each of the three goals (septic shock, SOFA score, survival) within minutes. The system is capable of predicting the most important EHR input characteristics of each of the three subjects. He compared the feature ranking results obtained using machine learning with those obtained using conventional univariate biostatistics coefficients [2].

Paheli Bhattacharya et al. proposed the unsupervised generalization algorithm DELSumm, designed to systematically incorporate the advice of legal experts into optimization tuning. An unsupervised algorithm that systematically integrates domain knowledge to extract summaries of court case documents. Approaches are measured for ROUGE1 and ROUGE2 scores using domain-specific independent supervised and unsupervised learning approaches [3].

Isabel Cachola proposed a learning system and strategy for applying a pre-trained language model that exploits the similarity between TLDR generation and the task of generating corresponding extreme sums and headers that exceed the baselines of robust extraction and abstract sums [4].

Hiroaki HagaID, Hidenori Sato et al. nine machine learning models were tested for all variants of the entire HCV genome sequence and the SVM accuracy ratio was 95.4% (kappa statistic, 0.90, *F* values, 0.94). This predictive model depends on the DAA used and should be considered for every new DAA going forward. The nine machine learning models compared for prediction accuracy are SVM, NN, KNN, LR, RF, and FDAGBM [5].

Gyanendra Chaubeyet al. described the use of three classification machine learning algorithms to classify people with reduced thyroid disease using a thyroid disease database: logistic regression classification, decision tree classification, and nearest neighbor classification. The authors detail data preparation, training, and testing, a step-by-step description of each method used, and a comparison of the accuracy of the methods used for prediction [6].

Pradeepika Verma and Anshul Verma discuss a reference model that combines the cohesive properties and coherent relationships of texts and use rhetorical and argumentative structures to reinforce referents based on lexical chains. The issues discussed are general and applicable to all possible scenarios when summarizing the text. We then discuss conventional extraction generalization methods that focus on the identified problem [7].

Hardik Gunjal et al. used convolutional neural networks were used to classify and represent complex features in medical record summaries using the MT sampling data set. He also uses GloVE to create pre-trained models for summarizing and classifying clinical statements. Convolutional neural networks were used to classify and represent complex features in medical record summaries using the MT sampling data set. He also uses GloVE to create pre-trained models for summarizing and classifying clinical statements [8].

Devansh Shah et al. present models based on supervised learning algorithms such as Naive Bayes, decision tree, nearest neighbor, and random forest, and various properties related to heart disease. It uses existing data sets from the Cleveland database of the UCI repository of cardiac patients. The data set has 303 instances and 76 properties. Of these 76 properties, only 14 are considered for testing, which is important to verify the performance of various algorithms. The results show that the highest accuracy score was achieved in the nearest neighbor [9].

Alexander R. et al. proposed an end-to-end model that integrates existing data extraction models with standard SDS models and provides competitive results for MDS data sets. The HiMAP model proposed by this author is similar to PGMMR in informational and fluency, but far superior in terms of lack of redundancy [10].

Mike Lewis et al. present a denoising autoencoder for pre-training a sequence model. BART is trained by training a model that (1) distorts text with a random denoising function and (2) reconstructs the original text. It uses, despite its simplicity, a standard transformer-based neural machine translation architecture that can be viewed as a generalization of BERT (due to bidirectional encoders), GPT (with left-to-right decoders), and other modern preconditioning plans [11].

Jessica López Espejel, convolution and recurrent neural networks are used to develop their original automatic synthesis methods for medical conversations between patients and healthcare professionals based on recent achievements [12].

Benjamin Stark et al. thirteen studies were reviewed from six different databases. These studies can be divided into two categories: (i) machine learning and data mining and (ii) ontology-based and rule-based approaches. The study was summarized and evaluated in several dimensions: disease, data storage, interfaces, data collection, data preparation, platforms/technology, and algorithms [13].

Reeta Rani et al. present literature reviews including abstracts and abstract summaries of texts. This overview paper provides an overview of various past research and research in the field of automatic text summarization [14].

Liang Yao et al. proposed a new way to classify the clinical text classification that combines the engineers of the management function, including the disease name, alternate name, and negative or infinite words that contain trigger syntax, and then use this trigger syntax to predict classes very restricted, and finally teach. Model CNN Word Built-in and UMLS CUIS Entity. The author identified this model with obesity data, and the evaluation result for obesity issues shows this method. This method exceeds the method of OTTUS for the operation. In this study, CNN model is to study effective hidden functions and CUIS attachments are useful for building clinical text expressions [15].

Saiyed aziya begum et al. proposed an evaluation method that is divided into external evaluation according to the degree of contribution to the completion of a specific task and internal evaluation, which directly evaluates the quality of the resume without being tied to a specific task. The performance of the automatic summary score can be measured using precision, recall, and F-score. Precision is the number of sentences found in both the system summary and the ideal summary divided by the number of sentences in the system summary. Recall that the number of sentences found in both the system and perfect summaries is divided by the number of sentences found in perfect summaries. F-score is a combined score of combining precision and recall. A metric set called Recall-Oriented Repetition of Essence Score (ROUGE) was introduced and became the standard for automated resume scoring, with human-generated model summaries and machine summaries [16].

A summary of the literature survey has been tabulated in Table 1.

3   Methodology

Medical transcription is a doctor–patient contact document as well as a current medical examination of a patient. This patient examination is very secure and confidential. Hiding the identity of a patient these texts are freely distributed, thus studying and analyzing, extracting meaningful information for understanding, interpreting, and interpreting is very complex and difficult. This may include identifying facts, gaps, and links between medical terminology data, speculation, or evidence. A complete analysis requires a final evaluation, from which a summary and prediction framework are extracted.

This section explains:

Table 1 Literature summary

Summary of literature					
S. No.	Author's name	Source	Year	Content/dimension/variable	Finding
1	Xiangming Wang et al	Hindawi	2021	Random forest	Disease prediction model
2	Davide Chicco and Luca Oneto	Springer-BMC	2021	Random forest feature selection	Case study—ranking of sepsis patient record
3	Paheli Bhattacharya et al	ICAIL	2021	DELSumm	Unsupervised domain-specific approach with ROUGE metrics measure
4	Iqbal H. Sarker	Springer	2021	Integration of machine learning and IoT	Integration of ML in real-world applications
5	Isabel Cachola	DeepAI	2020	TLDR, BART, T5 transformer, GPT-2 transformer	Single-sentence summaries of scientific papers
6	Hiroaki Haga ID et al	PLOS ONE	2020	Comparing predictive accuracy of nine ML algorithms SVM, NN, KNN, LR, RF, FDAGBM, DT, NB	Most accurate learning method was the support vector machine (SVM) algorithm (validation accuracy, 0.95; kappa statistic, 0.90; F-value, 0.94)
7	Gyanendra Chaubey et al	Springer	2020	Logistic regression and decision tree	Prediction of thyroid disease
8	Pradeepika Verma and Anshul Verma	Science Direct Computer and Information Science	2020	Graph-based methods-G-FLOW	Graph-based summarization by using clustering and MMR approaches
9	Hardik Gunjal et al	ResearchGate	2020	CNN, GloVE	Summarized and classified clinical discharge summaries
10	Devansh Shah et al	Springer	2020	KNN, Random forest, Naïve Bayes	Heart disease prediction
11	Asim Sohail	Multidisciplinary review journal	2020	POS tagging	Detailed information about text summarization
12	Alexander R. Fabbri et al	Sci-direct	2019	HiMAP model PG-MMR	Documents concatenation and summarizing longer input documents

(continued)

Table 1 (continued)

Summary of literature

13	Mike Lewis et al	Association for Computational Linguistics	2019	SquAD, MNLI, ELI5, Xsum BART	Map corrupted documents to the original
14	Puja Gupta et al	Elsevier	2020	Deep learning-artificial neural network (DL-ANN)	The deep neural network produced higher accuracy of 78% and precision, recall, and F-measure value of 83.58, 81. 25, and 80%
15	Jessica López Espejel	BMC	2019	NLP approach	Review of different approaches of text summarization
16	Benjamin Stark, Constanze Knahl, Mert Aydin, Karim Elish	IJACSA	2019	Review study	Review study on medicine recommender system
17	Shahadat Uddin et al	BMC Medical Informatics and Decision Making	2019	Comparison study	Comparing different ML algorithm on diabetics data set-SVM—high accuracy
18	Reeta Rani and Sawal Tandon	IJCAR	2018	KNN in text classification and clustering	Automatic text summarization by using NLP and machine learning
19	Liang Yao, Chengsheng Mao, Yuan Luo	Springer-BMC	2018	Convolutional neural network	Rule-based feature engineering and knowledge-guided DL techniques for disease classification
20	Darcie A P Delzell et al	PubMed	2019	Random forest	Elastic net and support vector machine applied on CT scans of lung nodules from 200 patients
21	Ambedkar Kanapala et al	Springer	2017	SALOMON, HAUSS, LetSum, Decision express	Abstractive summarization of legal document

1. Database details and their source of information.
2. A framework for steps to improve our proposed model for predicting a specialized medical category based on factors derived from medical records. In summary of the text, this enormous database collected is categorized.
3. Another proposed model is based on machine learning in a categorized and abstract document prediction treatment based on keywords/keyword search with training accuracy, validation accuracy, F1 points, accuracy, and recall.

3.1 Medical Transcript Summarization (MTS) via NLP

Healthcare organizations generate a lot of text data in the form of medical transcript results. This data is organized or organized in the form of EHR fields. By organizing and compiling a patient's basic medical history, patient summary algorithms can make better communication and care, especially for patients with chronic illnesses, whose medical records usually contain hundreds of notes [1]. Text summarization methods are developed by modern NLP applications using in-depth advanced reading methods. Each language has a set of rules that are used during the development of these sentences, and these set rules are also known as grammar. Comprehension involves work such as mapping input provided in the native language and issuing useful presentations that analyze various aspects of language while natural language production and the process of producing logical phrases and sentences in the natural language involves text editing, sentence editing, and text fulfillment.

Tokenizer. Token generation is the process of breaking a wire into tokens. Token generation involves three steps. (a) Splitting complex sentences into words. (b) Understand the meaning of each word in relation to the sentence and ultimately reveal the meaning of the structure in the input sentence. (c) Labels for any number of consecutive Ngram words.

Stemming. Stemming algorithms work by taking into account a list of common prefixes and suffixes that can be found in a word and trimming the end or beginning of a word. Now this promiscuous reduction can sometimes succeed, but not always, so we guarantee that this method provides certain limitations.

Lemmatization. A lemma is to analyze the form of a word. This requires a detailed dictionary in which the algorithm can find associations of lemmas and forms. One of the most important things to consider is that headword output is the correct term. This differs from stemming when the output is called as part.

Abbreviations. There are several words in the English language such as I, Not, take, if, etc. which are very useful in the construction of a sentence, and without them the sentence would be meaningless but these words are not helpful. In natural language processing, this list of words also known as Font Names and NLTK in Python has its own shortcut list and they use the same by importing it from NTLK.corpus.

Speech Parts (POS). In English sentences, verb forms, noun, adjective, adjective, etc. Figure 1 shows how a word works in meaning and grammar within a sentence, a sentence can be more than one part of speech based on the context in which it is used. In POS marking, words are classified and marked according to parts of speech. The associated tags are listed with their part-of-speech counterparts in Table 2 [17].

TFIDF. Term frequency inverse-document frequency the TFIDF representation reflects the reputation of the words of document collection as a separate document. The successful search for the engine can be developed based on TFIDF rating potential that can represent known words of text. Related queries of the document set the relevance of the document. However, the frequency indicator of the estimates (IDF) is a vocabulary, so it prevents the applications of TFIDF indicators that dynamically change.
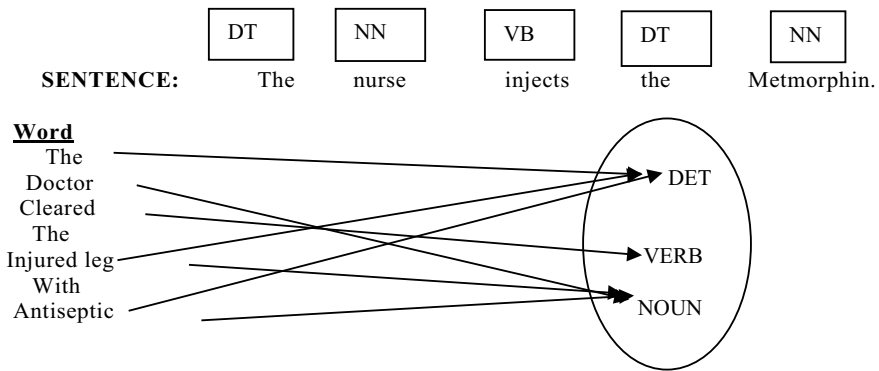
Fig. 1 Demonstration of POS Tag

Table 2 POS-tags and description

Tag	Description	Tag	Description
CC	Coordinating conjunction	PRP	Personal pronoun
CD	Cardinal number	PRPS	Possessive pronoun
DT	Determiner	RD	Adverb
EX	Existential there	RBR	Adverb, comparative
FW	Foreign word	RP	Particle
IN	Preposition or subordinating conjunction	SYM	Symbol
JJ	Adjective	To	To
JJR	Adjective, comparative	UH	Interjection
JJS	Adjective, superlative	VBG	Verb ground present participant
LS	List item marker	VBN	Verb, past participle
NN	Noun	VBZ	Verb 3rd person singular present
NNS	Noun, singular, or mass	PDT	Pre determiner

GloVe. The GloVe score [2] is the frequency with which words match other words. GloVe trains vectors after computing matches using dimensionality reduction. Other advantages of GloVe are its parallelization implementation and ease of training on large corpora. Embedding Word is a set of different language methods for modeling and learning features in the field of NLP.

Word2Vec. Word2Vec [3] creates distributed semantic representation of a word in a document. This model can be trained in the context of each word, so has a similar numerical view. Word2Vec is a predictive model to study vectors to reduce the loss of the target word in the context of this word.

Sentence to Vec. Sentence to vector is a word instead of words, Word2Vec presentation expansion. The vector representation of all the words at skip vector [18] issued in 2015 has good progress in the essence level attachment.

DOC2VEC. DOC2VEC [4] is an extension of Word2Vec or an extension of sentence to Vec. Single level embedding is the most commonly used for the investment of the word and is still a syntax and semantics on the word at an important meeting in the non-directional information only is limited to capture text. Context-dependent attachments, such as models such as ELM and LANGUAGE Bert, accounted for different levels.

ELMO. Insertion (ELMO)—vectors (ELMO)—vectors that vectors are used to prepare a vector using large houses to extract multilayer word of the attached file (BILM) model. ELMO studies a conceptual Word the view that captures the syntax, meaning, and word of a specific definition name of the ambiguous business (WSD).

BERT. The two-way rules of the encoder in the transformer (BERT) are based on both directional ideas of ELMO, but use the transformer [7]. BERT has been trained in advance to study the conditioning context of bidirections for all layers. Long-term vectors are available in complex NLP operations and can achieve results by the components of shading with only one additional output layer.

The methods that exist within the field of electronic health summary summaries have always taken and are indicative, meaning that abbreviations identify important parts of the original text rather than replacing the original text entirely. Few methods have been used in practice, and even a few have shown an impact on care quality and outcomes (Pivovarov 2016).

Text summary is defined in two ways: (1) unexplained summaries often require a combination of information, text congestion, and composition and requires an in-depth analysis of input and concept documents in text construction. (2) Summary released, most researchers focus on this summary because it incorporates key elements in other units (sentence, role) of documents. Sentence extensions are based on the very high end of the final summary. Summary strategies ranging from the issuance of an appropriate summary to the acquisition of sufficient observational data present a promising opportunity to use the study machine to predict the outcome and assign accurate treatment.
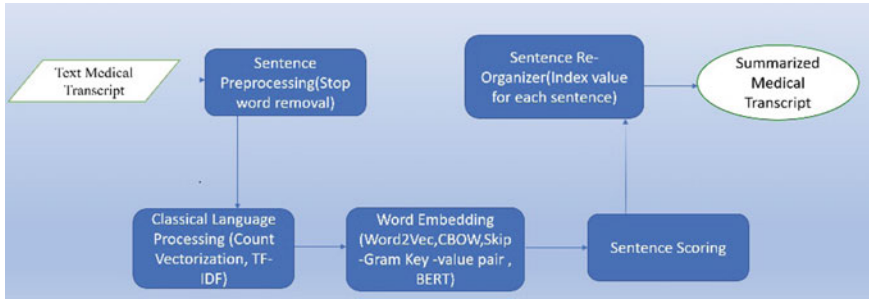
Fig. 2 Medical text summarization

Figure 2 shows the method for the automatic recording of medical notes from the medical transcript sample database [4]. We collect the MT sample data from the Kaggle website and continue to clean and process prior to this. 5000 patient records are taken in a nutshell, and this data is converted into our standard summary format for further analysis of treatment predictions. It includes the following outstanding features:

1. Template (Type): Specifies the type of text as a summary of patient discharge, diagnostic report, progress report, etc.
2. Disease (Name): The medical name or name of the problem the patient is being treated for.
3. Patient Information: Personal information such as age, gender, marriage, and smoker (Y/N).
4. Symptoms: a brief description of the patient's condition during the first visit.
5. Main Complaint: A specific issue that the patient is suffering from.
6. Medical History: A detailed description of the current illness, the patient's past years.
7. Allergies: An allergic patient has it.
8. Past Medical History: Definition of another patient's disease and prescribed treatment.
9. History of acquired illness: Information of any family members of an infected patient who may have it.
10. Key Indicators: Description of Essentials such as temperature and sugar.
11. Medical Diagnosis: Details of patient-centered diagnostic tests.
12. Testing: Other tests or tests performed or given to a patient such as admission to the ICU, glucose, or other treatment that need immediate attention.
13. Treatment: The definition of treatment given to a patient such as surgery and concrete.
14. Medications: It contains details of the drug and the dose given to the patient.
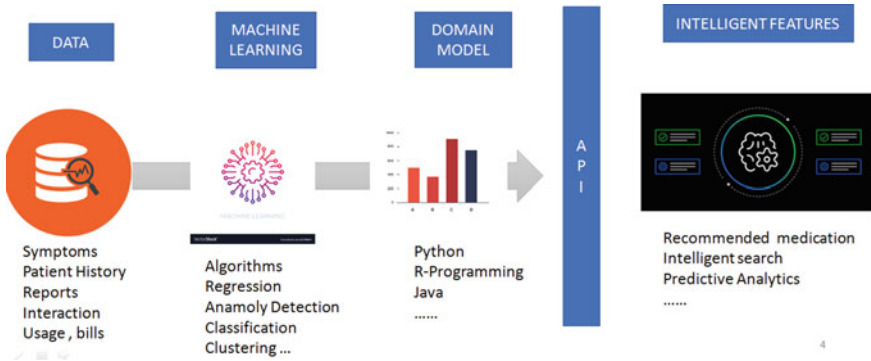
Fig. 3 Machine learning healthcare model

3.2 Machine Learning by Predicting Treatment

There are four main pillars to quality healthcare real-time patient monitoring, in-patient care, improved treatment options, and predictive diagnostics. All four pillars of quality health care can be effectively managed by predictable and technological methods. Figure 3 illustrates the machine learning model for healthcare industry.

Machine learning allows IT systems to recognize patterns and generate sufficient solution concepts based on existing algorithms and data sets. Machine learning uses mathematical methods to learn data sets. There are two main systems: the figurative method and the sub-symbolic method. A symbolic system, such as a suggestive system, in which the content of information is explicitly presented, i.e., established rules and models, but in reality, the system has less symbolic sensory networks. They affect the systems of the human brain in which the content of information is articulated.

In Fig. 4 a proposed model for medical prognosis, where a machine learning algorithm is applied to two models: one for prediction and one for interpretation. Predictive models are used to predict treatment using only summary medical records and are aimed at increasing accuracy. The second model is a legal delimiter used to describe only the results of the first model without worrying about accuracy. Machine learning is a method to obtain a complete combination of mathematical models from many explanatory variables (feature multiplication) based on knowledge of treatment outcomes and similar outcomes. Machine learning is difficult to understand compared to traditional mathematical methods, but it is very useful for big data analysis because it automatically detects combinations of different technological states [2].

Data Classification. First, we split the immature data into training data and test data according to section 14 professional selection. Selection and factor analysis has a significant impact on the development of training models. Some databases are useful, others are useless. Choosing obsolete features leads to choosing machine learning and less accurate models. Therefore, it uses mathematical and visual methods to analyze
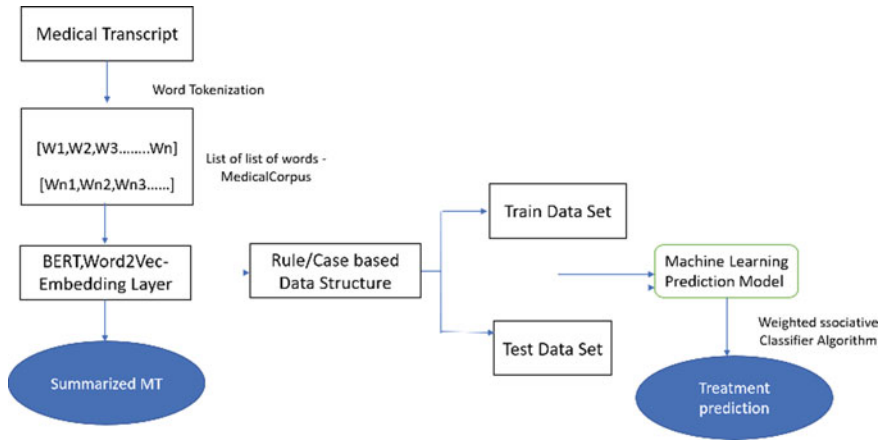
Fig. 4 Medical prognosis model

each factor, finds the most influential factors in the predicted results, and then selects the most powerful combination of factors based on the analytical learning results.

Training Model. Modeling is a repetitive process, a combination of different features, machine learning algorithms, and parameters to find the most suitable model.

Predictability: Processed test data is sent to the best model trained for forecasting.

4 Results

This paper presents a systematic literature search of medicine recommendation engines. We reviewed 21 studies, which can be divided into two categories: (i) machine learning and data mining-based, and (ii) rule-based approaches. Studies were summarized and assessed across multiple parameters: disease, data storage, interface, data collection, data preparation, platform/technology, an algorithm, and future work. Most studies not focused on disease included less information on data storage, interfaces, data collection, data preparation, platforms and technologies, and customized algorithms. For future work, our review proposes extending existing solutions by adding drug dosing recommendations and creating scalable solutions.

5 Conclusion

The work that needs to be integrated with research must overcome the need to consider based on the needs of certain medical technology. The integration of the technology is summarized in the prediction model of the training machine for precise treatment

preparation. Health statistics can reduce medical expenses, predict outcomes, prevent preventable diseases, and improve overall health. Life expectancy is increasing worldwide, which poses new challenges to modern medicine. All of this can help hospitals and doctors treat the right patients in the right way at the right time. For patients, timely assistance can be provided in certain emergency situations, and the data in this medical guide can help you find the right doctors and hospitals as a second way, reducing the amount of mistreatment and improving the quality of life. Health statistics can reduce medical expenses, predict outcomes, prevent preventable diseases, and improve overall diseases.

References

1. Wang X, Dong B (2021) Smart medical prediction for guidance: a mechanism study of machine learning. J Healthc Eng 2(6)
2. Chicco D, Oneto L (2021) Data analytics and clinical feature ranking of medical records of patients with Sepsi, no 14. BMC-Springer, p 12
3. Bhattacharya P et al (2021) Incorporating domain knowledge for extractive summarization of legal case documents. In: Eighteenth international conference on artificial intelligence and law, pp 22–31
4. Cachola I (2020) TLDR: extreme summarization of scientific documents. Deep AI
5. Hiroaki Haga ID, Sato H, Koseki A et al (2020) A machine learning-based treatment prediction model using whole genome variants of hepatitis C virus. PLoS One (8)
6. Chaubey G, Bisen D, Arjaria S et al (2020) Thyroid disease prediction using machine learning approaches. J Healthc Eng (44)
7. Verma P, Verma A (2020) A review on text summarization techniques. J Sci Res 64(1)
8. Gunjal H et al (2020) Text summarization and classification of clinical discharge summaries using deep learning
9. Shah D, Patel S et al (2020) Heart disease prediction using machine learning techniques, no 345. Springer, Berlin
10. Fabbri AR, Li I, She T et al (2019) Multi-news: a large-scale multi-document summarization dataset and abstractive hierarchical model. In: Proceedings of the 57th annual meeting of the Association for Computational Linguistics, pp 1074–1084
11. Lewis M, Liu Y et al. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension
12. Espejel JL (2019) Automatic summarization of medical conversations, a review. Project: Automatic summarization in medical domain
13. Stark B et al (2019) A literature review on medicine recommender systems. (IJACSA) Int J Adv Comput Sci Appl 10(8)
14. Rani R, Tandon S (2018) Literature review on automatic text summarization. Int J Curr Adv Res 7. ISSN: O: 2319-6475, ISSN: P: 2319-6505
15. Yao L, Mao C, Luo Y (2018) Clinical text classification with rule-based features and knowledge-guided convolutional neural network. In: Sixth IEEE international conference on healthcare informatics (ICHI 2018), pp 4–7
16. Saziyabegum S (2016) Smart medical prediction for guidance: a mechanism study of machine learning. Int J Comput Appl 156(12):28–36
17. Gaurav K, Kumar P (2017) Consumer satisfaction rating system using sentiment analysis. In: Smart cities, innovation, and sustainability, vol 10595. ISBN: 978-3-319-68556-4
18. Sarker IH (2021) Machine learning: algorithms, real-world applications, and research directions, vol 160. Springer, Berlin

19. Sohail A (2020) Methodologies and techniques for text summarization: a survey. Multi Rev J 7(11)
20. Garg S, Gupta P (2020) Breast cancer prediction using varying parameters of machine learning models. Procedia Comput Sci 171:593–601
21. Uddin S et al (2019) Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inf Decis Making 19–281
22. Peter T, Delzell DAP, Smith M et al (2019) Machine learning and feature selection methods for disease classification with application to lung cancer screening image data. Front Oncol 9:1393
23. Kanapala A (2019) Text summarization from legal documents: a survey. Springer, Berlin, pp 371–402
24. Jiang F et al (2017) Artificial intelligence in healthcare: past, present, and future. 2(4):230–243. PMID: 29507784; PubMed Central PMCID: PMC5829945
25. Hall MK, Pare JR, Venkatesh AK et al (2015) Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach, vol 23. PubMed, pp 269–278
26. Hung JC, Lin KC, Zhang KY et al (2016) Feature selection based on an improved cat swarm optimization algorithm for big data classification. J Supercomput 72:3210–3221
27. Bei Y, Xing W (2016) Medical health big data classification based on KNN classification algorithm. IEEE Access 8:28808–28819
28. Li Y, Jiang C (2019) Medical health big data classification based on KNN classification algorithm. IEEE Access 7:176782–176789
29. Shah NH, Callahan A (2017) Machine learning in healthcare. In: Key advances in clinical informatics. Academic Press, Cambridge, pp 279–291
30. Hu H, Wu X, Luo B, Tao C, Xu C, Wu W, Chen Z (2018) Playing 20-question game with policy-based reinforcement learning. arXiv preprint arXiv:1808.07645
31. Li J, Cheng K, Wang S, Morstatter F, Trevino RP, Tang J, Liu H. Feature selection: a data

A Screening Model for the Prediction of Early Onset Parkinson's Disease from Speech Features

Amisha Rathore and A. K. Ilavarasi

Abstract Parkinson's disease is a progressive neurodegenerative condition that is characterized by a variety of motor and non-motor disorders. Lack of movement, uncontrolled shaking of the hands, jaw, arms, legs, tongue, or hands are examples of motor symptoms. Non-motor symptoms include low blood pressure, sleep, pain, fatigue, skin and sweating, or restless legs. Clinical studies and assessment of associated symptoms, along with the characterization of a range of motor symptoms, are often used to diagnose Parkinson's disease. Traditional diagnostic procedures, on the other hand, may be vulnerable to diagnostic bias because they based on the assessment of motions that are usually delicate to human sight and hence hard to make early predictions. Patient therapy is complicated by the varied nature of Parkinson's disease symptoms and there is a variability in the progress. These symptoms are frequently missed, making early Parkinson's disease diagnosis is difficult. There are inadequate comparative data to suggest a specific treatment course in either early or advanced PD with motor difficulties. As a result, machine learning technologies have been used to overcome these issues and there is a great deal of interest in building models that recommend the treatment for Parkinson's disease. A preliminary study is conducted to characterize the features of importance. We then employ machine learning models to the speech dataset in order to derive insights on the diagnosis. The main objectives of this work are to propose the stack ensemble algorithm as a promising screening model for the early onset PD diagnosis. The algorithm exhibits 97% accuracy when compared to other models under evaluation.

Keywords Parkinson's disease · Machine learning algorithms · Diagnosis · Ensemble learning

A. Rathore
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

A. K. Ilavarasi (✉)
School of Computer Science and Engineering/Centre for Healthcare Advancement, Innovation and Research, Vellore Institute of Technology, Chennai, India
e-mail: ilavarasi.ak@vit.ac.in

1 Introduction

According to the Global Burden of Disease research 2018, the total prevalence of Parkinson's disease has almost quadrupled over the last two decades, from 2.5 quarter of patients in 1990 to 6.1 quarter of patients in 2016 [1]. According to estimates from 2016, India has roughly 0.58 million people living with Parkinson's disease, with a significant rise in prevalence projected in the following years. Given the immense number of individuals diagnosed by Parkinson's disease, there are little understanding into the fundamental environmental and genetic factors that really are specific to the Indian population.

Parkinson's disease symptoms usually begin gradually and gets worse over time. As the disease progresses, people with this condition may have trouble walking and communicating. Psychological and cognitive difficulties, as well as sleep disorders, tiredness, memory problems, and sadness, may occur. Both men and women can be affected by Parkinson's disease. Men, on the other hand, are more impacted by the illness than women. Parkinson's disease is strongly linked to a person's age. Despite the fact that the vast majority of people with Parkinson's disease acquire the condition beyond the year of 60, 5–10% of people with Parkinson's disease develop it before the range of 50 years of age. Parkinson's disease is usually, sometimes not, passed down from generation to generation, and some types have been related to particular genetic variants. Whenever nerve cells in the basal ganglia, a part of the brain that controls movement, become damaged or die, Parkinson's disease develops. Those nerve cells, normally create dopamine, an important brain neurotransmitter. When neurons become damaged, they release less dopamine, resulting in Parkinson's movement issues. Scientists are still unsure what causes dopamine-producing cells to die [2].

The Parkinson disease guidelines for pharmacists compiled the most important data and recommendations for pharmacists in four areas: diagnosis, communication, neuroprotection, prognosis, and treatment of both motor and non-motor symptoms of Parkinson's disease. Medications can help with concerns like walking, mobility, and tremor management [2]. These medications either increase dopamine levels or function as a dopamine substitute. Treatment suggestions and guidelines have been supplied by individual authors, as well as national and international associations and professional bodies. Some are evidence-based, such as the recently published national guidelines from Germany and Sweden, while others are based on expert consensus. On a number of aspects, the guidelines disagree, and many questions remain unanswered. There are insufficient comparative data for many treatment alternatives to allow for evidence-based prescriptions of the only option. As a result, it becomes a very fascinating topic for a lot of interest in developing models that prescribe Parkinson's disease treatment.

The remaining of this work is arranged in the following manner. Section 2 presents related work on Parkinson's disease using a machine learning technique. The methodology is described in depth in Sect. 3. Section 4 summarizes the results. Finally, Sect. 5 provides the conclusion and future prospects.

2 Literature Survey

Parkinson's disease is a degenerative neurological ailment including both motor and non-motor symptoms that affects all elements of action, especially plan, beginning, and execution. Previous to cognitive behavioral abnormalities, such as dementia, movement-related signs such as tremor, rigidity, and difficulty in starting might be noted. PD has a significant impact on patients' life quality, community activities, and family relations, as well as imposing significant financial costs on individuals and society. Traditionally, motor symptoms have been used to diagnose Parkinson's disease. Regardless of the fact that fundamental indications of Parkinson's disease have already been identified in patient assessment, most of the assessment scale used to determine symptom severity have still not been well reviewed and verified.

Despite the reality that non-motor signs including concentration and planning challenges, sleeping problems, and other sensors abnormalities like smell difficulty are present in many people before they develop Parkinson's disease, this lost clarity, are difficult to evaluate, and depend on the circumstances. As a result, non-motor symptoms cannot currently be used to diagnose Parkinson's disease on their own, however, some have been considered as diagnostic criteria. Many people experience motor problems such as wearing-off and dyskinesias after years of dopaminergic medication, particularly levodopa. Although numerous pathophysiological explanations for motor fluctuations may exist, therapeutic techniques aimed at reducing pulsatile dopaminergic activation may be advantageous [3]. As a result, the majority of treatment strategies aim to postpone the onset of motor fluctuations for as long as possible. Although dopaminergic therapy might temporarily relieve motor symptoms, motor variations can occur even in the early stages of the disease. It is also crucial to remember that each patient's non-motor traits may performs a significant impact, and these must be considered before deciding on a device-assisted treatment. Non-motor qualities, on the other hand, should not be utilized to determine which therapy to give because few research have looked into their effects [3]. Most nerve cells of people with Parkinson's disease have clusters of an enzyme called alpha-synuclein. Researchers are working to understand pathological functions and alpha-normal synuclein's physiological, including its link to genetic defects that cause Parkinson's disease. Despite the fact that certain cases of Parkinson's disease appear to be inherited, and some may be linked to specific gene changes, this disease was observed to attack at randomness, and so does not appear to occur frequently in the lot of instances. According to several studies, Parkinson's infection is characterized by a mix of interaction between genetic and environmental variables, including exposure to chemicals [4]. A multitude of causes can induce Parkinson's clinical signs. Parkinson disease is a word used to denote persons who have Parkinson's-like characteristics but are not suffering from the disease. While these disorders can be mistaken for Parkinson's disease at first, clinical diagnostics and pharmacological therapy response can help to distinguish them. Even though many disorders have common characteristics but require different treatments, it's vital to acquire an early diagnosis as soon as possible [5–7]. There seem to be currently no laboratory or

blood tests that can be used to diagnose Parkinson's disease. A diagnosis is made using medical records and a mental state examination [8–10]. Other important aspect of Parkinson's disease is the ability to recover after commencing the treatment.

3 Methodology

3.1 Feature Extraction Methods

This selection becomes even more significant when the number of features is huge, and our data is large and contains numerous attributes, therefore, these approaches are used to convert raw data into meaningful information and select the best features and attributes while eliminating the redundant ones [11].

3.1.1 Correlation Matrix with Heatmap

Seaborn is a one of the Python packages for data visualization. It provides a way to present data in graph format that is both instructive and appealing to the eye. Analysis of the data exploratory requires the use of correlation matrices. Correlation heatmaps provide the same data in a more visually attractive form. Correlation heatmaps do not perform well for the standard colormap. A diverging color palette with drastically distinct colors at both ends of the value range as well as a pale, virtually colorless midpoint works even better.

3.1.2 Univariate Selection

The primary goal of univariate feature selection is to choose the best feature based on a statistical test. We compare all of the attributes to the target attribute and then choose the best one. The link between the feature and the target variable is then examined. Each feature has a score in this analysis.

3.1.3 ExtraTreesClassifier Method

ExtraTreesClassifier or extremely randomized trees classifier seems to be a form of ensemble learning technique that aggregates the result of multiple de-correlated decision trees yielding the classification output. It splits all the observation to guarantee that the model does not overfit the data. It is conceptually identical to a random forest classifier, with the exception of how the decision tree algorithm within forest is constructed.

3.2 Filter Methods

Filter techniques are much faster than wrapper methods since they do not require the models to be trained. Wrapper approaches, on the other hand, are computationally expensive as well. Filter techniques evaluate a subset of characteristics using statistical methods, whereas wrapper methods employ cross validation [12].

3.2.1 Variance Threshold

A simple baseline technique to feature selection is the variance threshold. It eliminates all features whose variance falls below a certain level. It eliminates all zero-variance traits by default, that is, characteristics which have the same value across all samples. We think that attributes with a larger variance contain more important information, but keep in mind that we are not accounting for the link between feature variables or between feature and goal variables, which is one of the downsides of filter approaches.

3.2.2 Mean Absolute Difference

The only difference between mean absolute difference and variance threshold is that Mean Absolute Difference(MAD) does not include a square. MAD is a formula that calculates the absolute difference between two values. The absence of a square in MAD measures is the main distinction between variance and MAD measures.

3.2.3 Information Gain

The reduction in entropy caused by altering a dataset is calculated as information gain. Information gain can also be utilized for feature selection by assessing the benefit from every variable in the perspective of the target variable. The calculation is referred to that as reciprocal information between the different random variables in this subtly different usage.

3.3 Embedded Methods

Embedded approaches generally continuous in that they look after every step of the model training phase, precisely extracting those attributes that contribute more to the training for that iteration. Embedded techniques are a hybrid of filter and wrapper methods that incorporate feature interactions while being computationally efficient. We employed two embedded approaches LASSO regularization, random forest importance in this work.

3.3.1 LASSO Regularization

Least absolute shrinkage and selection operator is shortened as LASSO. It is a combination of variable selection and regularization that happens all at once. It is nothing more than a hybrid of linear regression with L1 regularization. Regularization is a method of reducing coefficients to zero. When a coefficient is 0, the feature is not considered, and it is effectively eliminated.

3.3.2 Random Forest Importance

A random forest is a technique for aggregating a set of decision trees. We must rank the trees according to the purity of the node in a random forest tree-based technique. Start of the tree includes greatest decrease in impurity and the other end of the tree includes least decrease in the impurity.

4 Results

The Parkinson's Disease Dataset is from Kaggle dataset repository [13]. There are 196 rows and 24 columns in the Parkinson dataset. There are 23 columns in all, with one target attribute. We discovered with 22 columns that are relevant and irrelevant column name was removed during data preprocessing. We used a machine learning technique to find the most associated feature among them after preprocessing.

We employed the correlation matrix with heatmap, univariate selection, and Extra-TreesClassifier method approach in the feature selection process. After applying feature selection as shows in Table 1, it is determined that MDVP:FO (Hz) is correlated, as it is the result from two different feature-related algorithms.

Table 2 shows the applied filter methods. The variance threshold makes all output true, indicating that all features are connected to the target. The most comparable feature discovered by mean absolute error is MDVP:Fhi (Hz). As far as Information gain measure is concerned, it was discovered that PPE is the most closely connected trait.

As stated in Table 3, we also used various embedded methodologies, such as LASSO regularization (L1) and random forest importance for lasso regularization all features are correlated, while PPE for random forest is strongly correlated.

Table 1 Feature extraction methods

Algorithm	Resultant feature
Correlation matrix with heatmap	MDVP:FO (Hz)
Univariate selection	MDVP:Flo (Hz)
ExtraTreesClassifier method	MDVP:FO (Hz)

Table 2 Filter methods

Algorithm	Resultant feature
Variance threshold	ALL
Mean absolute difference (MAD)	MDVP:Fhi (Hz)
Information gain	PPE

Table 3 Embedded methods

Algorithm	Resultant feature
LASSO regularization (L1)	ALL
Random forest importance	PPE

Table 4 contains all of the machine learning algorithms we have looked at so far. And, because MDVP:FO (Hz) is the most related feature in both the correlation matrix with heatmap and ExtraTreesClassifier methods, and with information gain and random forest importance, PPE is the highly correlated feature for filter methods and embedded methods. So, we conclude that MDVP:FO (Hz) and PPE are the most associated features (Table 4).

Table 4 Comparison table

Algorithm	Resultant feature	Algorithm	Resultant feature
1	Feature extraction methods	Correlation matrix with heatmap	MDVP:FO (Hz)
		Univariate selection	MDVP:Flo (Hz)
		ExtraTreesClassifier method	MDVP:FO (Hz)
2	Filter methods	Variance threshold	ALL
		Mean absolute difference (MAD)	MDVP:Fhi (Hz)
		Information gain	PPE
3	Embedded methods	LASSO regularization (L1)	ALL
		Random forest importance	PPE

Table 5 Comparison table for classification algorithms

Algorithm	Accuracy with all features (%)	Accuracy with relevant features (%)
Logistic regression	94	79
Support vector machines	25	79
Extreme gradient boosting	97	97
Stacking ensemble	97	97

5 Conclusion and Future Work

The main purpose of this work is to develop therapy options based on a person's Parkinson's disease symptoms. We use machine learning techniques to identify the features that are strongly linked to the goal attribute. Model accuracy is investigated from the speech dataset which give significant inferences about early onset Parkinson's in the clinical perspective. Experiments on feature selection methods are conducted (Table 1), and result comparison is carried out with a wide band of machine learning algorithms (Table 5). Stack ensemble algorithm yields promising results in terms of accuracy and the model itself selects the best features from base learners mitigating the time complexity.

References

1. de Lau LML, Breteler MMB (2006) Epidemiology of Parkinson's disease. Lancet Neurol 5(6):525–535
2. Jankovic J (2008) Parkinson's disease: clinical features and diagnosis. J. Neurol Neurosurg Psychiatry 79(4):368–376
3. Patel T, Chang F (2015) Practice recommendations for Parkinson's disease: assessment and management by community pharmacists. Can Pharmacists J/Rev des Pharmaciens du Can 148(3):142–149
4. Olanow CW, Stern MB, Sethi K (2009) The scientific and clinical basis for the treatment of Parkinson disease (2009). Neurol, 72(21 Supplement 4):S1–S136
5. Chaudhuri K, Rizos A, Sethi KD (2013) Motor and nonmotor complications in Parkinson's disease: an argument for continuous drug delivery? J neural transm 120(9):1305–1320
6. Ahmadi Rastegar D, Ho N, Halliday GM, Dzamko N (2019) Parkinson's progression prediction using machine learning and serum cytokines. NPJ Parkinson's disease 5(1):1–8
7. Proulx CE, Beaulac M, David M, Deguire C, Haché C, Klug F, Kupnik M, Higgins J, Gagnon DH (2020) Review of the effects of soft robotic gloves for activity-based rehabilitation in individuals with reduced hand function and manual dexterity following a neurological event. J Rehabil Assistive Technol Eng 7, p. 2055668320918130
8. Kapsalyamov A, Hussain S, Sharipov A, Jamwal P (2019) Brain–computer interface and assist-as-needed model for upper limb robotic arm. Adv Mech Eng 11(9):1687814019875537
9. Dorsey ER, Elbaz A, Nichols E, Abbasi N, Abd-Allah F, Abdelalim A, Murray CJ (2018) Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 17(11):939–953
10. Palmerini L, Rocchi L, Mellone S, Valzania F, Chiari L (2011) Feature selection for accelerometer-based posture analysis in Parkinson's disease. IEEE Trans Inf Technol Biomed 15(3):481–490
11. Brewer BR, Pradhan S, Carvell G, Delitto A (2009) Feature selection for classification based on fine motor signs of Parkinson's disease. In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 214–217
12. Gunduz H (2021) An efficient dimensionality reduction method using filter-based feature selection and variational autoencoders on Parkinson's disease classification. Biomed Signal Process Control 66:102452
13. https://www.kaggle.com/datasets/vikasukani/parkinsons-disease-data-set

Speed Control of Hybrid Electric Vehicle by Using Moth Flame Optimization

Krishna Prasad Naik, Rosy Pradhan, and Santosh Kumar Majhi

Abstract The main objective of this paper is to control the throttle valve of the nonlinear hybrid electric vehicle by controlling the electronic throttle control system (ETCS). Optimization-based control is very popular in recent days. Moth flame optimization (MFO) is a metaheuristic optimization that has been used to solve many engineering problems. MFO exhibits considerable exploitation and exploration to find the global optimum solution. In this work, a PID controller that is tuned by MFO algorithm has been used to obtain the parameters such as K_p, K_i and K_d for controlling the throttle valve of Internal Combustion Engine (ICE). The proposed techniques perform better in terms of rise time, settling time and maximum overshoot for the system while comparing with Ziegler-Nichols and Hand-tuning method.

Keywords ETCS · PID controllers · SSA · MFO algorithm

1 Introduction

In the last few years, the mushrooming growth of vehicles leads to increase of emission resulting the global warming and heavy environmental pollution. In addition, the rise of cured oil price has become a global economic issue.

By looking these issues, researchers and scientist invented the concept of electric vehicles (EV) which is environmentally friendly. However, there is a big issue on charging the battery of EV. To overcome this issue, the hybrid electric vehicles (HEV) concept came to the market which is powered by a conventional internal

K. P. Naik (✉) · R. Pradhan
Department of Electrical Engineering, Veer Surendra Sai University of Technology, Burla, India
e-mail: kpnaik92@gmail.com

R. Pradhan
e-mail: rosypradhan_ee@vssut.ac.in

S. K. Majhi
Department of Computer Science and Engineering, Veer Surendra Sai University of Technology, Burla, India
e-mail: smajhi_cse@vssut.ac.in

combustion engine vehicle (ICEV) and an electric motor. This hybridized model ensures hazardous exhaust emission and fuel utilization reduction. HEV is used to utilize the electrical energy storage and uses for the drive systems. ICE provides the required mechanical power to drive the vehicle, and on the other hand, the energy used by HEVs is obtained from an engine and additional source of energy [1–5].

There are several HEV according to their build model like series, parallel, series–parallel, complex and plug-in HEVs. Generally, this is a complex system as it combines with mechanical and electrical elements which consist of interdisciplinary technology. Its performance is based on interrelated control factors.

Kumar Yadav et al. [6] presents the dynamic modeling, characteristics and behavior of HEV by using MATLAB for simulation purpose. The authors developed certain techniques like FLC to find out the parameters of PID controller, state feedback control methods like LQR, PPT, OBC, and ZN method and Hand-tuning method. Kumar et al. [7] explain the mathematical model of dynamic behavior for HEVs. The authors explained the correctness of FLC-based methods for PD and PI controllers in a cascade control loop and give an optimum result. Gaur et al. [8] described methods for the control, design and development of HEV. In this paper, GA is used to find out different parameters of the controller and suggested a new algorithm that gives optimum results. Borhan et al. [9] proposed an HEV model for energy management and power split based on model predictive control. This model can operate as both series and parallel HEVs.

It has been studied from the literature that different optimization techniques such as GWO, PSO, ACA, ALO have been applied to tune the PID and FOPID controllers [10–17]. In this exertion, MFO has been considered to tune the PID controller for HEV systems.

The remaining part of this paper is organized as follows: Section 2 explains about the model description of HEV and system behavior. Section 3 explains the analysis of the controller, Sect. 4 shows the results and discussion, and Sect. 5 contains the conclusion and future scope.

2 System Description

2.1 Model Description

In this work, the main contribution to the current research is to control or smoothen the speed of the HEVs. Figure 1 shows a throttle valve that represents the inlet point of the ICE. Where ICE is controlled with DC servo motor. There is a bar/throttle plate whose position is controlled by an electronic throttle control system (ETCS) which is responsible for the achievable wide range of vehicle speed. Flow of air enters into the manifold of ICE. The main aim of the bar in the throttle valve is to control the flow of air. So that a controller is required to control the highly nonlinear ETCS [6–9]. There are several controllers are available like PID controller, FLC and state feedback controllers such as LQR, OBC and PPT. In this work, used a PID controller

to control the ETCS. Many techniques are developed to find out the parameters of the PID controller such as Hand-tuning method, ZN method and recently developed some metaheuristics algorithms also.

Figure 2 represents the architecture of HEV in a closed-loop control system. There is a reference input (R_d), output (y), error signal (e), controller, plant, disturbance (K_g), and the controlling line represents the flow of information. The performance of the HEVs depends on load of the vehicle, weight of the vehicle and various road grades, i.e., slope of road, surface of road.

In accordance with Newton's second law, variation of the vehicle speed can be predicted as a function of the applied force. The dynamic equation of the vehicle is specified as:

$$\omega \frac{ds}{dt} = K e(\theta) - \sigma s^2 - K_g \tag{1}$$
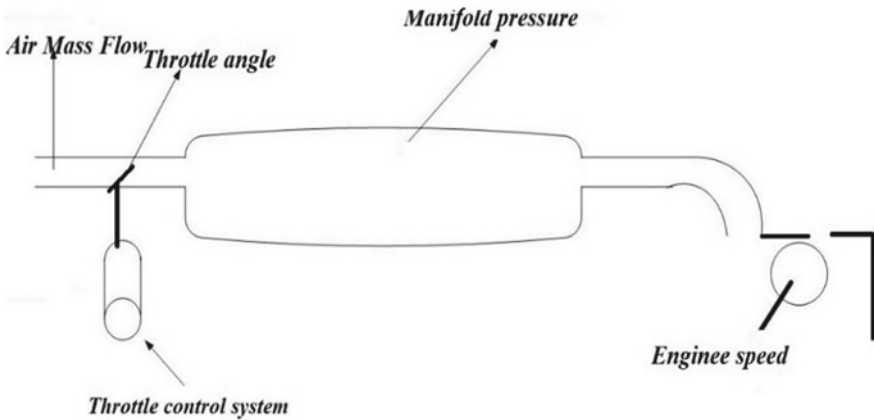
Fig. 1 Schematic diagram of ETCS of ICE
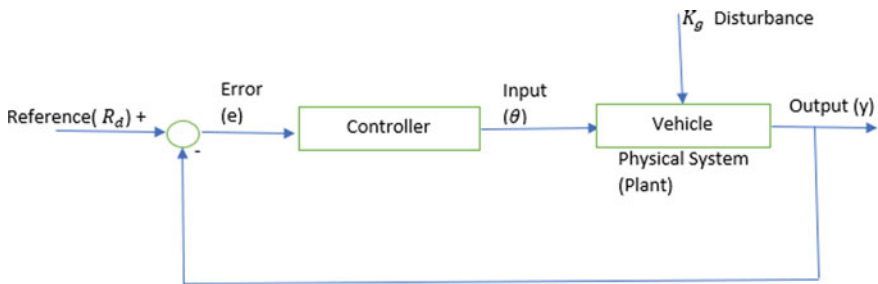
Fig. 2 Electric vehicle control system

Table 1 Numerical values of HEV

Constant	Values (SI unit)
Idle force of the engine (K_i)	6400 N
Time constant of the engine (δ_e)	0.2 s
Coefficient of the engine force (γ)	12,500 N
Mass of the vehicle (ω)	1000 kg
Drag coefficient (σ)	4 N/(m/s)^2

$$\delta_e \frac{dK_e(\theta)}{dt} = -K_e(\theta) + K_1(\theta) \tag{2}$$

$$K_1(\theta) = K_i + \gamma\sqrt{\theta} \tag{3}$$

where

K_e a position of the throttle valve of ICE
K_g gravity-induced force, road grade function
s vehicle speed
θ throttle position
ω mass of vehicle and
δ_e time constant of engine.

Some assumptions are required to control the speed:

- Engine is very fast as compared to the vehicle.
- The vehicle will mostly be operated around some nominated velocity (s_0) more or less on a flat road and proceed to linearize the differential Eqs. 1–3.
- Gravity-induced force (K_g) is 30% weight of the vehicle.
- Engine time constant (δ_e) is very negligible and generally varies in between the range 0.1 and 1 s; here, 0.2 s is considered (Table 1).

2.2 Plant Model Simplification and Linearization

The state-space analysis method has been used to solve a nonlinear equation. The state-space equations are given in Eqs. (4) and (5).

$$x'(t) = Ax(t) + Bu(t) \tag{4}$$

$$y(t) = Cx(t) + Du(t) \tag{5}$$

where

A, B, C and D are the matrices, $y(t)$ is the output, $u(t)$ is the input, and $x(t)$ is a state vector.

Characteristics equation of the state space is $|SI - A| = 0$. The value of A, B, C, D is obtained by solving the dynamic Eqs. (1–3) of the vehicle by using the characteristics equation of state-space analysis method and utilizing the above-mentioned assumption.

$$A = \begin{bmatrix} 0 & 0.001 \\ 0 & -5 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 829,000 \end{bmatrix}, C = [1 \ 0], D = [0 \ 0] \qquad (6)$$

By solving the above matrix, the eigen values of the vehicle are $\lambda_1 = 0$, $\lambda_2 = -5$. Transfer function of the vehicle is specified in Eq. 7.

$$\frac{y(s)}{\delta(s)} = \frac{829,000}{s^2 + 5s}. \qquad (7)$$

2.3 PID Controller

Most widely used in industrial purpose is PID controller, which gives several advantages. It is used to decrease the steady state error and simultaneously increase the stability. It has pole at origin, and two zeros are there. One zero compensates the pole, and other is increase the stability. The transfer function of the Proportional–Integral–Derivative controller is

$$G_c(s) = K_p + K_i S + K_d S G_c(S) = K_p(1 + 1/T_i + T_d S) \qquad (8)$$

In this exertion, MFO algorithm is used to find out the different control parameters of PID controller such as K_p, K_i and K_d while setting an optimum result that we discuss on the later section. Simulation-based model of HEV with PID controller is shown in Fig. 3 [18].
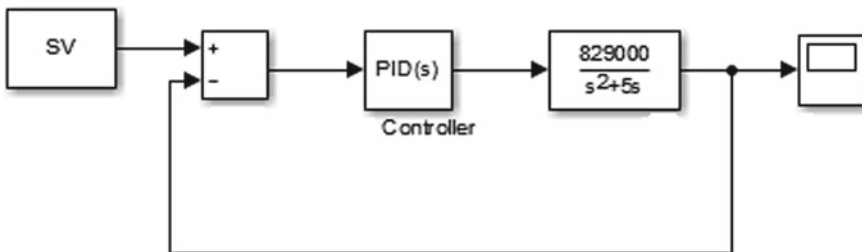
Fig. 3 Simulink model of HEV

3 Optimal Design of Proposed Controller

3.1 Moth Flame Optimization

From the study of literature, moths are delighted to the moonlight and fly linearly with a fixed angle. This navigation mechanism is called transverse orientation [10, 19, 20]. Moths are trapped by the artificial lights and move around these artificial lights in a spiral way because of their inability the transverse orientation. Moths are the beginning point, and the flame is the ending point. Population-based algorithms have two basic stages for a superior result that is exploitation, i.e., finest solution and exploration, i.e., unknown search space. MFO algorithm can optimize the local optima problem. Moths and flames are the candidate solutions. Their positions are their respective problem variables. First moth always upgrades its position w.r.t best flames and last moth upgrade its position w.r.t inferior flames. After updated moths, the flames are shorted based on their fitness values. Due to that, moths update their position w.r.t corresponding flames. Specific flames are provided to each moth to prevent local stability.

3.2 Proposed MFO-Based PID Controller

Here moths are stored in a matrix $(B_{n,v})$ and flames are stored in a matrix $(L_{n,v})$. Their fitness values are stored in an array form subsequently. Each moth has a position vector. Fitness function of moth is calculated based on their respective position. In this case, moths are search agents. Moths are navigating in the logarithmic spiral path around its flames. For each and every time, they find a satisfactory solution and the flame position is updated every time [10, 19].

Some notation and their description are: B is the moth's matrix, L is the flames matrix, OB and OL represent to store their fitness values of moth and flame, respectively.

$$B = \begin{bmatrix} B_{1,1} & B_{1,2} & \ldots & B_{1,v} \\ B_{2,1} & B_{2,2} & \ldots & B_{2,v} \\ \vdots & \vdots & \vdots & \vdots \\ B_{n,1} & B_{n,2} & \ldots & B_{n,v} \end{bmatrix}, L = \begin{bmatrix} L_{1,1} & L_{1,2} & \ldots & L_{1,v} \\ L_{2,1} & L_{2,2} & \ldots & L_{2,v} \\ \vdots & \vdots & \vdots & \vdots \\ L_{n,1} & L_{n,2} & \ldots & L_{n,v} \end{bmatrix} \tag{9}$$

n constitutes number of moths, v constitutes number of variables.

Subsequent fitness values of moths and flames are

$$OB = \begin{bmatrix} OB_1 \\ OB_2 \\ \vdots \\ OB_n \end{bmatrix}, OL = \begin{bmatrix} OL_1 \\ OL_2 \\ \vdots \\ OL_n \end{bmatrix} \tag{10}$$

MFO algorithm calculates the global optimization problem with three tuples.

$$MFO = (I, P, E) \tag{11}$$

Function I generates the random population of moths and follows their fitness values. Function P represents moths travel in search space, and Function E represents the termination/ending the criterion.

$$B_i = S(B_i, L_j) \tag{12}$$

where B_i is ith moth and L_j is jth flame, S is the spiral function. Moth's matrix B is determined by

$$B(i, j) = (ub(i) - lb(i)) * rand() + lb(i) \tag{13}$$

ub and lb are upper and lower boundaries, respectively.

The mathematical equation of logarithmic spiral path of moths is in Eq. 14

$$S(B_i, L_j) = G_i.e^{bt}.\cos(2\pi t) + L_j \tag{14}$$

where "b" is constant value, which is used for determining the logarithmic shape, and "t" is a random number $[-1$ to $1]$. G_i represents the distance in between ith moth and jth flame.

$$G_i = |L_j - B_i| \tag{15}$$

The number of flames during each iteration is defined as

$$R = round\left(Q - h \times \frac{\text{current iteration}}{\text{maximum iteration}} \right) \tag{16}$$

where Q is the max. number of flames and "h" is current iterations.

3.3 Estimation of PID Controller

In this work, PID controller is tuned with the MFO. An objective function called Integral Absolute Error (IAE) is used to minimize the parameters of the PID controller

to obtain optimum results [8].

$$F = \int_0^\infty |e(t)|\mathrm{d}t \qquad (17)$$

Equation 17 represents the considered IAE objective function (Fig. 4).
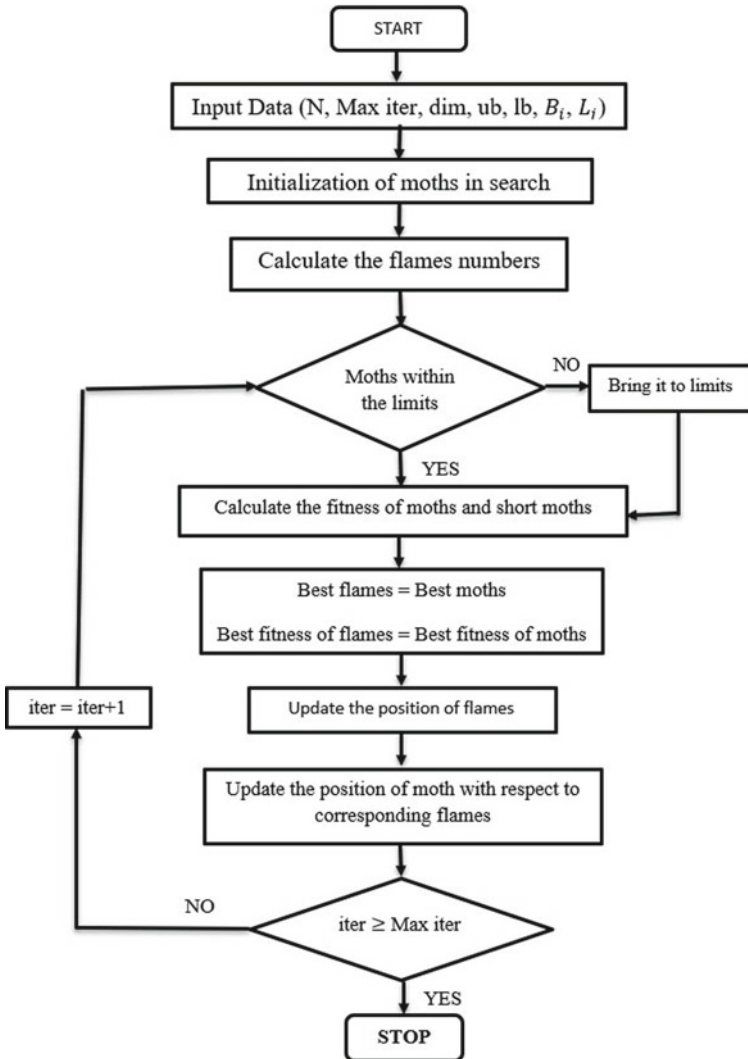
Fig. 4 Flowchart of MFO-PID controller

4 Results and Discussion

This section describes the response and analysis of HEVs with open-loop and closed-loop system. The comparative study to get the optimum responses is explained.

Figure 5 represents the open-loop response for HEVs.

Figure 6 shows performance and response of the PID controller using Hand-tuning rule and Ziegler-Nichol's approach.

In Fig. 6, PID controller is tuned with Ziegler-Nichol's method and Hand-tuning rule which shows 46.7% and 15.6% of maximum overshoot, respectively. From the response, it can be concluded that better results can be obtained by using Hand-tuning rule [6].

Figure 7 shows that PID controller is tuned by MFO algorithm-based response. In this approach, MFO algorithm is considered and found the different parameters of the PID controller. The maximum iteration is 700, and different upper and lower bounds are chosen to get the optimum results. MFO algorithm gives 8.15% maximum overshoot and comparatively better results with rise time and settling time.

Table 2 shows all parameters such as max. iteration, dimension, K_p, K_i and K_d of MFO-PID controller for HEV.

The best optimal value of the objective function found by MFO is 0.12064.

Table 3 represents the rise time, settling time, % of max overshoot, % of steady-state error of different methods.

From Fig. 8 and Table 3, it can be concluded that the optimum results are obtained by using moth flame optimization (MFO)-based PID controller. MFO-PID gives least maximum overshoot, steady-state error and settling time, i.e., 8.15%, 0.05% and 0.56 s, respectively. After settling once, the rise time of MFO-PID is 0.11 s which is a very minimal disturbance. Consequently, MFO-PID gives optimal results when compared to ZN-PID and PID with Hand-tuning rule.
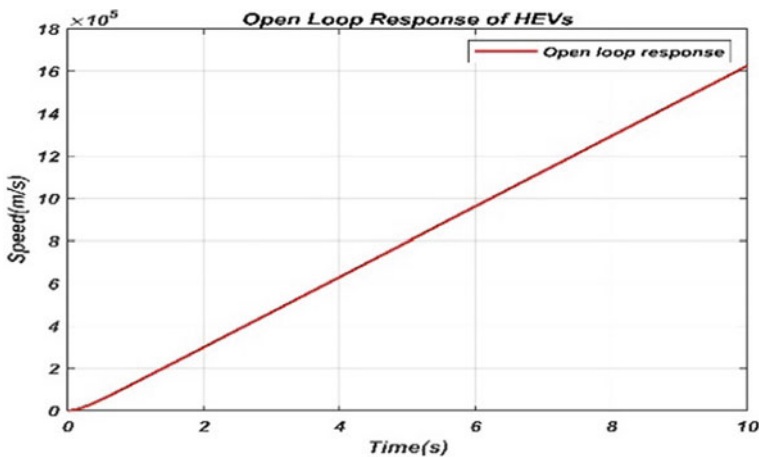
Fig. 5 Simulink response of open loop for HEV
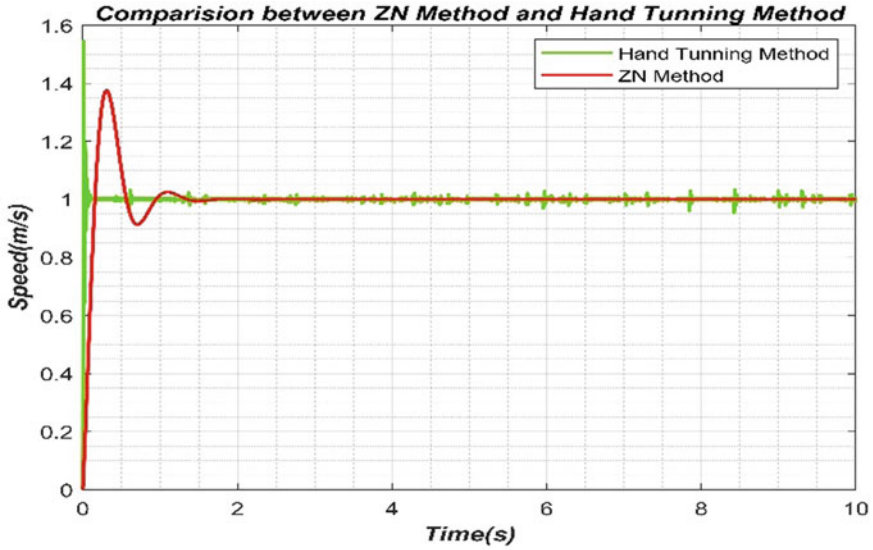
Fig. 6 Simulink response of ZN method and hand-tuning method
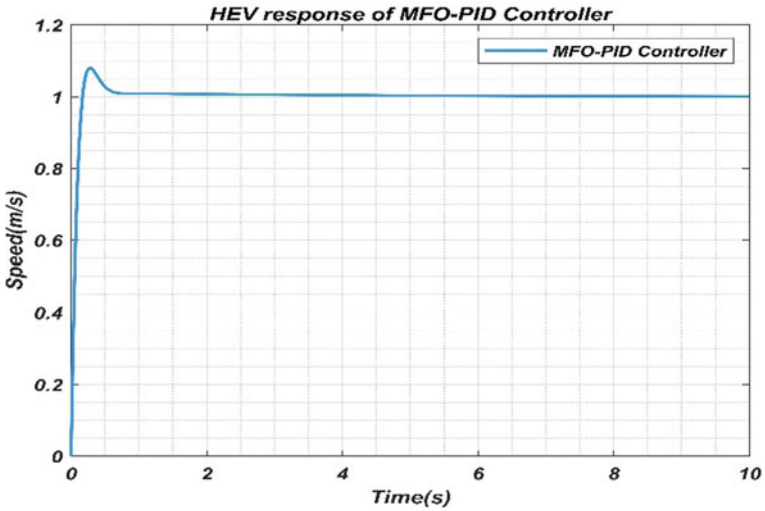
Fig. 7 Simulink response of MFO-PID controller

5 Conclusion

MFO-PID controller gives better performance for nonlinear HEV systems in comparison to other traditional control strategies. The proposed controller smoothens the

Table 2 Parameters of MFO-PID controller

Parameters	Values (SI unit)
Maximum iteration	700
Dimensions	3
K_p	1.3×10^{-4}
K_i	3.5×10^{-5}
K_d	1.39×10^{-5}

Table 3 Performance index of all control method for unit step input

Controllers	% Max overshoot	Rise time (s)	Settling time (s)	% Steady-state error
PID with ZN method	46.7	0.121	1.18	0.98
PID with hand-tuning rule	15.6	0.39	0.79	0.15
PID tuned with MFO algorithm	8.15	0.11	0.56	0.05
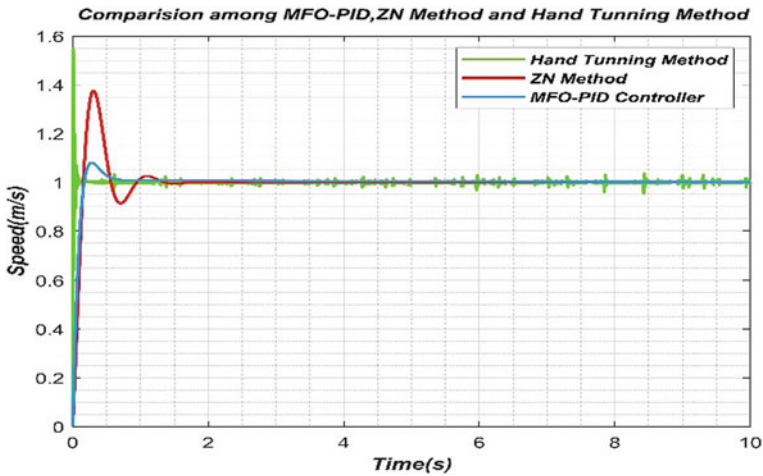
Fig. 8 Simulink response of MFO-PID, ZN method, hand-tuning rules

throttle position control and provides a wide range of speed based on various road grade. The maximum overshoot and settling time obtained by this controller are to achieve least current and torque. Due to this, it will improve the battery performance of the vehicle. It can be concluded from the above simulation results and performance index that the transient and steady-state response of MFO-PID controller gives optimal performance.

References

1. Khooban MH, Gheisarnejad M, Vafamand N, Boudjadar J (2019) Electric vehicle power propulsion system control based on time-varying fractional calculus: implementation and experimental results. IEEE Trans Intell Veh 4(2):255–264. https://doi.org/10.1109/TIV.2019.2904415

2. Assadian F, Fekri S, Hancock M (2012) Hybrid electric vehicles challenges: strategies for advanced engine speed control [Online]. Available: http://ieeexplore.ieee.org

3. Enang W, Bannister C (2017) Robust proportional ECMS control of a parallel hybrid electric vehicle. Proc Inst Mech Eng Part D J Automobile Eng 231(1):99–119. https://doi.org/10.1177/0954407016659198

4. Perng JW, Lai YH (2016) Robust longitudinal speed control of hybrid electric vehicles with a two-degree-of-freedom fuzzy logic controller. Energies 9(4). http://doi.org/10.3390/en9040290

5. Luo C, Shen Z, Evangelou S, Xiong G, Wang FY (2019) The combination of two control strategies for series hybrid electric vehicles. IEEE/CAA J Automatica Sinica 6(2):596–608. https://doi.org/10.1109/JAS.2019.1911420

6. Kumar Yadav A, Gaur P, Kant Jha S, Gupta JRP, Mittal AP (2011) Optimal speed control of hybrid electric vehicles. J Power Electron 11(4):393–400. https://doi.org/10.6113/JPE.2011.11.4.393

7. Kumar V, Rana KPS, Mishra P (2016) Robust speed control of hybrid electric vehicle using fractional order fuzzy PD and PI controllers in cascade control loop. J Franklin Inst 353(8):1713–1741. https://doi.org/10.1016/j.jfranklin.2016.02.018

8. Kaur J, Saxena P, Gaur P (2013) Genetic algorithm based speed control of hybrid electric vehicle. In: Sixth international conference on contemporary computing (IC3). IEEE, pp 65–69. https://doi.org/10.1109/IC3.2013.6612163

9. Borhan H, Vahidi A, Phillips AM, Kaung ML, Kolmanovsky IV, Cairano SD (2012) MPC-based energy management of a power split hybrid electric vehicle. IEEE Trans Control Syst Technol 20(3):593–603

10. Mirjalili S (2015) Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl-Based Syst 89:228–249. https://doi.org/10.1016/j.knosys.2015.07.006

11. Pradhan R, Majhi SK, Pradhan JK, Pati BB (2018) Antlion optimizer tuned PID controller based on Bode ideal transfer function for automobile cruise control system. J Ind Inf Integr 9:45–52

12. Pradhan R, Majhi SK, Pradhan JK, Pati BB (2020) Optimal fractional order PID controller design using Ant Lion optimizer. Ain Shams Eng J 11(2):281–291

13. Pradhan R, Majhi SK, Pradhan JK, Pati BB (2017) Performance evaluation of PID controller for an automobile cruise control system using ant lion optimizer. Eng J 21(5):347–361

14. Pradhan R, Majhi SK, Pradhan JK, Pati BB (2021) Oppositional Crow Search algorithm with mutation operator for global optimization and application in designing FOPID controller. Evol Syst 12(2):463–488

15. Pradhan R, Majhi SK, Pradhan JK, Pati BB (2019) Design and performance evaluation of fractional order PID controller for heat flow system using particle swarm optimization. In: Computational intelligence in data mining. Springer, Singapore, pp 261–271

16. Pradhan R, Majhi SK, Pradhan JK, Pati BB (2019) Design of fractional order PID controller for heat flow system using hybrid particle swarm optimization and gravitational search algorithm. Int J Comput Intell Stud 8(1–2):59–72

17. Pradhan R, Majhi SK, Pradhan JK, Pati BB (2019) Comparative performance evaluation of fractional order PID controller for heat flow system using evolutionary algorithms. Int J Appl Metaheuristic Comput (IJAMC) 10(4):68–90
18. Bisht P, Yadav J (2020) Optimal speed control of hybrid electric vehicle using GWO based fuzzy-PID controller. In: Proceedings—2020 international conference on advances in computing, communication and materials, ICACCM 2020, pp 115–120. http://doi.org/10.1109/ICACCM50413.2020.9212985
19. Shah YA, Habib HA, Aadil F, Khan MF, Maqsood M, Nawaz T (2018) CAMONET: Moth-Flame Optimization (MFO) based clustering algorithm for VANETs. IEEE Access 6:48611–48624. https://doi.org/10.1109/ACCESS.2018.2868118
20. Sayed GI, Hassanien AE (2018) A hybrid SA-MFO algorithm for function optimization and engineering design problems. Complex Intell Syst 4(3):195–212. https://doi.org/10.1007/s40747-018-0066-z

A Novel Hybrid Algorithm for Effective Quality of Service Using Fog Computing

Rajendar Janga, B. Kumara Swamy, D. Uma Vishveshwar, and Swathi Agarwal

Abstract A novel hybrid algorithm for effective quality of service using fog computing is implemented. In recent years, there has been an increasing interest in solving the over-provisioning and under-provisioning of elastic cloud resources because of the Service-Level Agreement violation problem. The recent studies have reported that fog cloud services may serve as a better elastic cloud model over a single provider model. A major problem with the federated cloud is the interoperability between multiple cloud service providers. Therefore, novel hybrid algorithm for effective quality of service using fog computing is introduced. In this initially, data is passed to the cloud layer. Next from cloud layer to fog computing layer. Fog computing layer arranges its data in parallel form and passes the data to the fog device. To manage the fog device hybrid algorithm is utilized based on the memory and CPU. In cloud data centers resources are utilized if fog device is not available. Now data is computed based on virtualization. The virtualized data is visible to the user device. At last from results it can observe that novel hybrid algorithm for effective quality of service using fog computing improves the security, efficiency and provides services in effective way.

Keywords Fog computing · Cloud layer · Fog device · Data processing · Fog computing layer

R. Janga (✉) · B. Kumara Swamy · D. Uma Vishveshwar
Department of CSE-Data Science, CMR Engineering College, Kandlakoya, Hyderabad, Telangana 501401, India
e-mail: janga.rajendar@gmail.com

B. Kumara Swamy
e-mail: kumaraswamy.b@cmrec.ac.in

S. Agarwal
Department of CSE, CVR College of Engineering, Vastunagar, Hyderabad, Telangana 501510, India

1 Introduction

Fog computing brings out various discernments in various individuals, for example, for a few, it alludes to getting to programming and putting away information in the cloud portrayal of the Internet or a system and utilizing related administrations. For other people, it appears as the same old thing, however only a modernization of the time-sharing model that was generally utilized in the 1960s before the appearance of generally lower-cost computing stages [1]. Fog computing is here, and there saw as a resurrection of the great centralized computer customer worker model. Fog computing empowers buyers to get to assets online through the web from anyplace whenever without stressing over specialized and physical administration or upkeep issues of the first assets. Besides, fog computing assets are dynamic and adaptable. Fog computing is autonomous computing and is absolutely not normal for the matrix and utility computing. Google Apps is the central case of cloud computing; it empowers to get to administrations by means of the program and could be conveyed on a great many machines over the Internet [2].

Assets are open from the cloud whenever and from wherever over the world utilizing the web. Fog computing is less expensive than other computing models. The upkeep cost included is just about zero since the specialist co-op is answerable for the accessibility of administrations and customers are liberated from support and the executive's issues of the asset machines. Because of this element, fog is otherwise called utility computing or basically IT on request.

Adaptability is key quality of fog computing and is accomplished through worker virtualization. Fog computing gives calculation, programming, information access and capacity benefits that do not require end-client information on the physical area and setup of the framework that conveys the administrations.

Fog computing is one of the most remarkable developments that has gotten extravagant of technologists around the world. While fog computing has gigantic focal points, for example, adaptability, fast versatility, estimated administrations and most significant of them the potential that it has for cost reserve funds to the undertakings, it likewise has a lot of security chances that no undertaking can bear to ignore. The security dangers exude from the wide scope of the weaknesses intrinsic in a fog computing framework, and without dependable security orders, there is an obvious hesitance with respect to associations to receive a generally an incredible domain called cloud computing.

Security and hazard appraisal would incorporate investigation of the effect of assortment of dangers and assaults on different parts of cloud computing including; transformation of cloud computing, upkeep of mystery and security of individual information, access and refreshing of information [3].

The connection between IoT devices and cloud datacenters is nothing but distributed computing. The extended type of distributed computing is nothing but fog computing. Based on the devices service of computing and storage is provided. Mainly it consists of switches, routers, base stations and networking components. There are own respective services of storage, computing and networking for the

above components. This concept is introduced by the Cisco by limiting the cloud extensions [4]. Fog computing is mainly used in the applications of health care, real time and gaming applications.

Consequently, the distinguishing proof of most suitable arrangement orders to reinforce security and protection in the cloud condition has gotten vital to all business activities in the cloud. The subject of the exploration study 'Security Threats and Attacks on fog Computing System: An Empirical Study' is not just very important what's more, contemporary yet additionally a fascinating test to improve confirmation level what's more, confidence of associations by dependably alleviating security dangers to diminish the security dangers in this new space of fog computing [5].

In this investigation, which are investigating and breaking down the noticeable information security and system security assaults on the cloud frameworks. Existing investigations uncover that DoS (DDoS, XDoS, HDoS) assault and man-in-the-center assault are more unmistakable assaults on the cloud systems. Likewise, malware infusion assaults with two classes to be specific SQL infusion and Cross Site Scripting (XSS) assaults are generally normal and unmistakable information security assaults on the part of cloud systems. Our investigation stays limited on DDoS assaults and plans to give the counter action calculations and propose the answers for information protection from malware infusion assaults [6]. Further, our undertaking is to draw out the purposes behind hesitance for appropriation of cloud toward accomplishing the third goal of the study.

2   Security and Privacy Issues in Fog Computing

Trust
Basically for computing applications, services of security provided are highly reliable. Therefore, in fog network trust is provided at certain level. In the networks of IoT and fog, authentication plays major role. But this authentication service level is not enough to overcome the attacks of malicious [7]. Therefore, for the purpose of fostering relations trust plays very significant role while interacting. Hence in fog network trust plays two way roles. The main intent of fog nodes is to validate the services based on the request that provide for the data in genuine way.

Authentication
In the fog network one of the most important requirements to provide better security is services provided should be very effective. In the fog network the data should become the part of network to provide authentication and to access the services. If the data not becomes the part of network, then authentication is not provided for the data which cause attacks. In various ways the challenges are involved which will process the storage.

Secure Communications in Fog Computing
In this fog nodes are offloaded based on the requirements of storage and processing. Requirements of security will be minimum but there should be implementation of

IoT devices [8]. Hence in communications of IoT devices security care should be taken. Because of this there will be high security provided to the devices and attacks will be reduced. Once data is processed based on offload then only fog nodes interact with IoT devices.

End User's Privacy
On the distribute nodes, fog computing provides the security. This will reduce the data center pressure by providing end to end privacy for data. Sensitive data is collected by the fog computing.

Malicious Attacks
There are various types of malicious attacks obtained in the environment of fog computing. But they do not depend on the security measures. This will determine the network capabilities [9]. Denial of Service (DoS) is one of the types of malicious attack. The devices are the not authenticated mutually which are not connected to device. Hence it attacks directly. Hence in the storage devices, devices should be connected to the authentication to avoid from malicious attacks [10].

3 Novel Hybrid Algorithm Using Fog Computing

Figure 1 shows the flowchart of novel hybrid algorithm using fog computing. In this initially, data is passed to the cloud layer. Next from cloud layer to fog computing layer. Fog computing layer arranges its data in parallel form and passes the data to the fog device. To manage the fog device hybrid algorithm is utilized based on the memory and CPU. In cloud data centers resources are utilized if fog device is not available. Now data is computed based on virtualization. The virtualized data is visible to the user device.

Algorithm
Step-1: In this initially, data is passed to the cloud layer.

Step-2: Fog computing layer arranges its data in parallel form and passes the data to the Fog device.

Step-3: To manage the fog device hybrid algorithm is utilized based on the memory and CPU.

Step-4: In cloud data centers resources are utilized if fog device is not available.

Step-5: Now data is computed based on virtualization.

Step-6: The virtualized data is visible to the user device.

Table 1 shows the comparison table of fog computing and novel hybrid algorithm using fog computing. In this cost, security, complexity and efficiency parameters are used. Compared with fog computing, novel hybrid algorithm using fog computing reduces the cost and complexity and increases the security and efficiency.

Figure 2 shows the comparison of cost and security for fog computing and novel hybrid algorithm using fog computing. Compared with fog computing, novel hybrid algorithm using fog computing reduces the cost and increases the security.
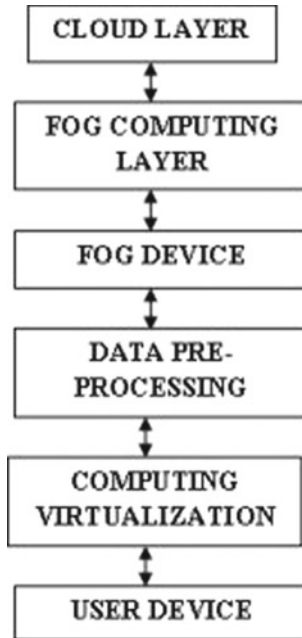
Fig. 1 Flowchart of novel hybrid algorithm using fog computing

Table 1 Comparison table

S. No.	Parameter	Novel hybrid algorithm using fog computing (%)	Fog computing (%)
1	Cost	12	84
2	Security	95	38
3	Complexity	9	31
4	Efficiency	97	27
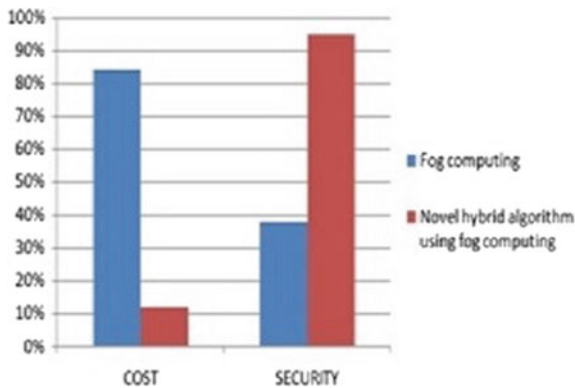
Fig. 2 Comparison of cost and security for fog computing and novel hybrid algorithm
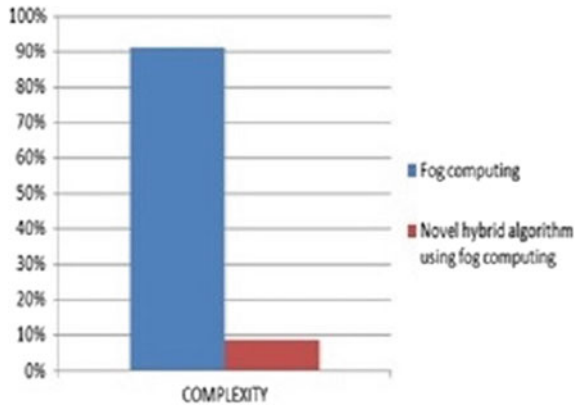
Fig. 3 Comparison of complexity for fog computing and novel hybrid algorithm
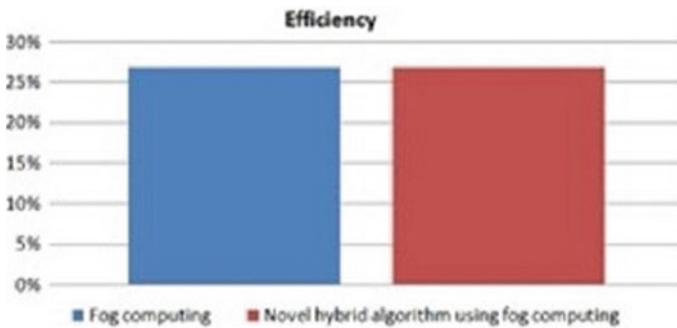
Fig. 4 Comparison of efficiency for fog computing and novel hybrid algorithm

Figure 3 shows the comparison of complexity for fog computing and novel hybrid algorithm using fog computing. Compared with fog computing, novel hybrid algorithm using fog computing reduces the complexity in effective way.

Figure 4 shows the comparison of efficiency for fog computing and novel hybrid algorithm using fog computing. Compared with fog computing, novel hybrid algorithm using fog computing increases the complexity in effective way.

4 Conclusion

Hence in this paper a novel hybrid algorithm for effective quality of service using fog computing was implemented. The main intent of this paper is to improve the security and reduce the complexity based on the possibilities of cloud computing. Efficiency and quality of system are improved based on the applications of development. At

last from results it can observe that novel hybrid algorithm for effective quality of service using fog computing improves the security, efficiency and provides services in effective way.

References

1. Skarlat O, Bachmann K, Schulte S (2018) Fogframe: IoT service deployment and execution in the fog. KuVS-Fachgespräch Fog Comput 1:5
2. Naha RK, Garg S, Georgakopoulos D, Jayaraman PP, Gao L, Xiang Y, Ranjan R (2018) Fog computing: survey of trends, architectures, requirements, and research directions. IEEE Access 6:47980–48009
3. Abderrahim M, Ouzzif M, Guillouard K, Francois J, Lebre A (2017) A holistic monitoring service for fog/edge infrastructures: a foresight study. In: The IEEE 5th international conference on future internet of things and cloud (FiCloud 2017)
4. Tsai P-H, Hong H-J, Cheng A-C, Hsu C-H (2017) Distributed analytics in fog computing platforms using tensorflow and kubernetes. In: 2017 19th Asia-Pacific network operations and management symposium (APNOMS). IEEE, pp 145–150
5. Gupta H, Vahid Dastjerdi A, Ghosh SK, Buyya R (2017) iFogSim: a toolkit for modeling and simulation of resource management techniques in the internet of things, edge and fog computing environments. Softw Pract Experience 47(9):1275–1296
6. Brogi A, Forti S (2017) Qos-aware deployment of IoT applications through the fog. IEEE Internet Things J 4(5):1185–1192
7. Delicato FC, Pires PF, Batista T (2017) The resource management challenge in IoT. In: Resource management for internet of things. Springer, Berlin, pp 7–18
8. Perera C, Qin Y, Estrella JC, Reiff-Marganiec S, Vasilakos AV (2017) Fog computing for sustainable smart cities: a survey. ACM Comput Surv (CSUR) 50(3):32
9. Kaur K, Dhand T, Kumar N, Zeadally S (2017) Container-as-a-service at the edge: trade-off between energy efficiency and service availability at fog nano data centers. IEEE Wirel Commun 24(3):48–56
10. Povedano-Molina J, Lopez-Vega JM, Lopez-Soler JM, Corradi A, Foschini L (2013) DARGOS: a highly adaptable and scalable monitoring architecture for multi-tenant clouds. Futur Gener Comput Syst 29(8):2041–2056

Hierarchical Learning of Outliers

Gouranga Duari and Rajeev Kumar

Abstract Outlier analysis and data clustering both are important tasks in the data analysis domain. In this paper, we decompose data points at multiple levels of hierarchy to identify the presence of outliers at multiple levels. This approach creates a tree-like structure of outlier's hierarchically. Experimenting on a real dataset gives us vital information about the characteristics of outliers and numerical results establish strong evidence for validating the proposed approach.

Keywords Hierarchical clustering · Decomposition · Hierarchy · Outlier

1 Introduction

Detection of outliers remains an essential and extensive research area in data analytics due to its mass applications. Researchers can access necessary information from outliers, which helps in making actionable decisions about data patterns. The fundamental aim of clustering is to recognize homogeneous clusters from a set of data points. The prime goal of clustering is to divide the whole dataset into clusters such that data points in the same cluster are more similar to each other than to data points from other clusters. Clustering is often used as a primary step for data analytics in the unsupervised learning process.

Hierarchical clustering with outliers. Although k-means is well-known clustering algorithm in unsupervised data analytics, it often performs poorly on real-world datasets. This is because the k-means assumes that all of the data points can be naturally partitioned into k distinct clusters with the same size, which is often an unfeasible assumption in practical application. Real-world data typically have outliers, and the k-means method is highly sensitive to it. Outliers can dramatically alter the quality of the clustering solution, and it is very very important to consider this in

G. Duari (✉) · R. Kumar
Data to Knowledge (D2K) Lab, School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: gourangaduari5@gmail.com

designing algorithms for the detection of outliers. In the hierarchical clustering technique, the clusters do not forcefully include all the data points. So, a few excluded data points are likely candidates to be the outliers.

Our Contribution. Motivation to this work is from the contribution of Kumar and Rockett, who investigated a Learning-follows-Decomposition (LFD) strategy [1] for hierarchical learning of complex high-dimensional problems. In this paper, we propose the first algorithm for the hierarchies of outliers, which is a generalized approach for outliers in multiple-level of hierarchy. Here, we use hierarchical clustering to decompose the data points into multiple-level hierarchically so that we can separate the outliers from normal data points in each level of hierarchy. We use apriori knowledge of outliers in the datasets to spot the outliers in each level of hierarchy.

The rest of the paper is organized as follows. Section 2 provides the related work. Section 3 provides a brief discussion on the model. Section 4 describes the experimental process. Section 5 includes results and analysis. Finally, we conclude the paper in Sect. 6.

2 Related Work

Kumar and Rockett [1] proposed a generic solution for high-dimensional problems using hierarchical Learning. This Learning-follows-decomposition strategy decomposes a problem into a series of subproblems based on fitness for purpose, a set of function approximators are assigned to each subproblem so that each module learns to specialize in a subdomain.

Beyan and Fisher [2] proposed a new hierarchical decomposition approach for imbalanced datasets. The hierarchy is constructed using the similarity of labeled data subsets at each level of the hierarchy with different levels being built by different data and feature subsets. Clustering is used to partition the data while outlier detection is utilized to detect minority class samples.

Paulhiem and Meusel [3] proposed the attribute-wise learning for scoring outliers (ALSO) approach, which searches directly for patterns in the dataset instead of exploiting density. Here, the outlier detection problem splits into a set of supervised learning problems to identify the patterns. This algorithm returns both the patterns underlying the data, as well as estimators for the strength of those patterns in each attribute simultaneously. A predictive model is assigned for each attribute, which predicts the values of that attribute from the values of all other attributes and computes the deviations between the predictions and the actual values. They derive both weights for each attribute using the deviation, and a final outlier score is assigned using those weights. The weights separate the relevant attributes from the irrelevant ones and thus make the approach well suitable for identifying the outliers.

Guha et al. [4] proposed a hierarchical clustering-based model called CURE. In this method, a fixed number of data points are generated by choosing well-scattered data points from clusters, then a specified fraction of them are moved toward the

center of a cluster. Here, more than one representative data point per cluster helps CURE to adjust the geometry of non-spherical shapes prudently; it also helps to reduce the effects of outliers in a cluster. CURE is robust to outlier detection in large datasets. Karypis et al. [5] proposed a novel hierarchical clustering technique called CHAMELEON. In their work, the similarity of two clusters is measured based on a dynamic model. If the closeness and inter-connectivity between two clusters are highly relative to the internal inter-connectivity of clusters and closeness of items within clusters, then two clusters are merged using a dynamic model. Gagolewski [6] proposed hierarchical clustering linkage criterion called 'Genie'. Here, two clusters are linked based on a specified economic inequity measure (e.g., the Gini-index or Bonferroni-index), and the cluster sizes does not increase above a prescribed threshold.

3 Model

3.1 Overview

This approach consists of two major parts: (a) hierarchical clustering and (b) spotting outliers in the cluster using apriori knowledge of the dataset. The data decomposition strategy creates levels of outliers in each iteration hierarchically to identify the potential likely outliers.

3.2 Hierarchical Clustering

Clustering-based algorithms [7] generally describe the behavior of the data by creating clusters of homogeneous data instances. Clustering-based techniques are unsupervised since they do not require any prior knowledge about data. Outlier detection is noticeably different from the clustering process, as the main aim of the clustering process is to identify the clusters for data points, while outlier detection is concerned detecting outliers. Detection of outliers is an unsupervised task. So, the clusters with smaller sizes consist of significantly fewer data points than other clusters are considered outliers. The performance of clustering-based techniques is highly dependent on the effectiveness of the clustering algorithm in capturing the cluster structure of the inconsistent instances. Hierarchical clustering partitions the set of data points into groups of different levels and creates a tree-like structure. In our approach, hierarchical clustering partitions the outliers in different levels and creates multiple levels of outliers hierarchically.

3.3 Spotting Outliers

The hierarchical clustering separates the outliers from normal instances in multiple levels using decomposition strategy [1]. Here, we use apriori knowledge of the datasets to spot the outliers in different levels of hierarchies manually. We continue the process until we segregate every single outlier from normal data instances. With this approach, we identify the presence of outliers at different levels, which can further investigate for more research output.

4 Experiment

4.1 Problem Definition

To describe the algorithm, let $X = \{x_1, x_2, \ldots, x_n\}$ be a numerical dataset containing n data points, each of which is described by d numerical attributes. Here, a distance function is assumed over pairs of data points of X as: $d : X \times X \longrightarrow R$. The most common setting to consider is when the dataset X consists of points in the d-dimensional Euclidean space, and the distance function between pairs of points in X is defined as the Euclidean distance. Euclidean distance is represented as: $d(x_i, x_j) = \sqrt{\sum_{t=1}^{d} |x_{it} - x_{jt}|^2}$, where $x_i = (x_{i1}, x_{i2}, \ldots, x_{id})$ is the representation of x_i in R^d. Here, we have used Ward's linkage criterion as it produces better clusters hierarchy.

At first, we consider the whole dataset as one cluster. We decompose the entire dataset into two clusters using the hierarchical clustering technique. Then, we check the presence of outliers in the respective clusters using apriori knowledge of the dataset. There are three possible outcomes of hierarchical decomposition with the context of outliers: (a) clusters without outliers, (b) clusters only with outliers, and (c) clusters with both outliers and inliers. We do not consider further decomposition in the first two cases. If we find the outliers in the clusters embedded with inliers, we further decompose that particular cluster using hierarchical clustering for learning. We continue the process until we separate the outliers from normal instances. This process creates multiple-level of outliers hierarchically.

4.2 Dataset

We have used only one benchmark dataset due to the scarcity of the required dataset. SatImage[1] [8] dataset is obtained from ODDS repository. This dataset contains 5803 instances with 36 dimensions, and all the attributes are numerical in nature. Here,

---

[1] http://odds.cs.stonybrook.edu/satimage-2-dataset/.

1.2% (71) data samples are labeled as 'outlier', and the rest of the data samples are 'normal'. As our approach is concentrated only on numerical data points, we have used all the attributes for our experiment.

4.3 Experiment Setup

We execute all our experiments[2] on the Windows 11 operating system, and it contains 8 GB of RAM. In all the experiments, algorithms are executed on Jupiter notebook in Python programming language. We use *sci-kit learn* library for implementing hierarchical clustering. We have used *sci-kit learn* library for standardized the data using the normalization method to make our model computation easier.

5 Results and Analysis

In this section, the performance of our approach is evaluated on the benchmark dataset. At first, we decompose the whole datasets at multiple levels using hierarchical clustering, we identify the presence of the outliers in multiple levels using apriori knowledge of outliers.

We have identified a total of 13 levels of outliers in the SatImage dataset, which are explicitly visualized in Fig. 1. A tree-like structure is formed with the identified outliers in the respective levels hierarchically. In Fig. 1, we can notice that outliers are distributed in the two different clusters at level-2. Potential sixty outliers are categorically formed a separate cluster of outliers at level-3, and this particular cluster has no normal data points.

We can also see that cluster with 11 outliers in level-2 has required very deep decomposition to segregate the outliers from the normal data points. We have identified the rest of the 11 outliers using apriori knowledge of the dataset at different levels of hierarchy, called level-3, level-6, level-9, level-10, level-11, and level-12, respectively. Table 1 has given a brief presentation of the numbers of identified outliers at their respective levels.

---

[2] https://github.com/gourangaduari1995/Hierarchical-Learning-of-Outliers.
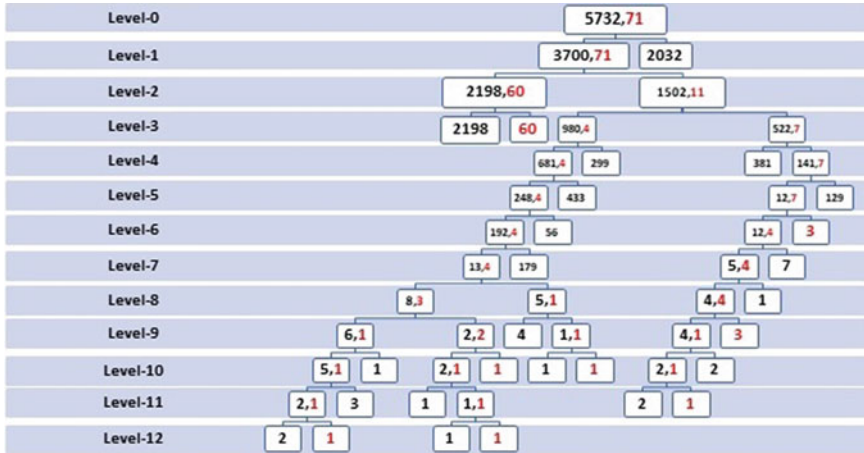
Fig. 1 Outliers are at multiple levels and they are presented with tree-like structure

Table 1 Outliers are categorically identified at multiple levels with a separate clusters of outliers only

Level-3	Level-6	Level-9	Level-10	Level-11	Level-12
60	3	3	2	1	2

Maximum identified outliers are shown bold

Here, we have presented a tree-like structure for the outliers at multiple levels hierarchically. We have presented each node with the number of outliers and inliers. Here, we have three types of nodes: (a) nodes with inliers and outliers, where corresponding numbers are highlighted by black and red color respectively; (b) nodes with inliers, where corresponding numbers are highlighted by black color; and (c) nodes with outliers, where corresponding numbers are highlighted by red color. We have started with level zero, and we have identified 13 such levels.

5.1 Discussion

We have given a detailed brief on results in the previous section. From the model-oriented results, we observe that three main categories of nodes emerge at each level with intrinsic information of outliers. We have identified 60 outliers at level-3, which is approximately 84.5% of the total outliers. These outliers are segregated very quickly with a few decompositions. So, these outliers are likely global outliers [9], which require a further level of investigation to understand detailed characteristics and behavior of the outliers at respective levels. We have decomposed further to understand the behavior of the rest of the 11 outliers, which are likely to be local outliers [9]. We have found that these outliers are densely embedded with the normal data points, so they require further investigation.

In this paper, we have identified the hierarchical position of outliers using decomposition strategy [1] to create the levels of outliers. These levels can be used to create the degree of outliers, which requires further investigation on how each level has inter-connections with the position of outliers and other related factors. These inter-connections can have a great role in deciding suitable detection techniques based on rationality.

The presence of noise [10] in the dataset is known to pose a problem, particularly for hierarchical clustering. In our approach, likely outliers are separated from normal data points using a decomposition strategy to create clusters of outliers caused by some systematic deviation of measurement, which may be affected by noise in the real dataset. So, the treatment of the noise factor is a matter of further investigation.

6   Conclusion

In this paper, we have proposed a hierarchy of outliers at multiple levels using a data decomposition strategy. This preprocessing strategy creates multiple levels of outliers, which can be used further for the detection of an outlier by choosing effective standard techniques at each level based on the characteristics and behavior of the respective levels. This approach can be robust in choosing a rational detection technique, which will essentially leverage on reduction of classification error rate.

References

1. Kumar R, Rockett P (1998) Multiobjective genetic algorithm partitioning for hierarchical learning of high-dimensional pattern spaces: a learning-follows-decomposition strategy. IEEE Trans Neural Netw 9(5):822–830
2. Beyan C, Fisher R (2015) Classifying imbalanced data sets using similarity based hierarchical decomposition. Pattern Recogn 48(5):1653–1672
3. Paulheim H, Meusel R (2015) A decomposition of the outlier detection problem into a set of supervised learning problems. Mach Learn 100(2):509–531
4. Guha S, Rastogi R, Shim K (1998) Cure: an efficient clustering algorithm for large databases. ACM Sigmod Rec 27(2):73–84
5. Karypis G, Han E, Kumar V (1999) A hierarchical clustering algorithm using dynamic modeling
6. Gagolewski M, Bartoszuk M, Cena A (2016) Genie: a new, fast, and outlier-resistant hierarchical clustering algorithm. Inf Sci 363:8–23
7. Aggarwal CC, Reddy CK (2014) Data clustering. Algorithms and applications. Data mining and knowledge discovery series. Chapman and Hall/CRC, London
8. Aggarwal CC, Sathe S (2015) Theoretical foundations and algorithms for outlier ensembles. ACM SIGKDD Explor Newsl 17(1):24–47
9. De Vries T, Chawla S, Houle ME (2010) Finding local anomalies in very high dimensional space. In: Proceedings of IEEE international conference on data mining. IEEE, pp 128–137
10. Cao J, Kwong S, Wang R (2012) A noise-detection based adaboost algorithm for mislabeled data. Pattern Recogn 45(12):4451–4465

Smart ECG Monitoring System Based on IoT

Bani Gandhi and N. S. Raghava

Abstract Presently, heart-related issues have become a main reason of apprehension for people around the globe. This is because of two reasons. Firstly, because of time constraint for reaching the hospital or providing a remedy, i.e., once the person gets an attack, an immediate action has to be taken to save the victim's life because it is just a matter of fraction of seconds. Secondly, the population is growing, even now in many parts of the world, proper infrastructure is missing, and in some areas, there is a shortage of clinics and hospitals. The scarcity of the staff in the healthcare domain has been a really big issue in the last decade globally. It has been anticipated that by 2030, approximately 23.6 million people will die from CVDs, and it will remain as a major cause of deterioration in the healthcare domain. The solution to the problem is provided by a concept that has taken an edge over conventional techniques used in the field of health care and is referred to as 'e-health'. e-health systems gather, process, and evaluate the medical data. The people who actually make use of these devices are either health conscious people, people who are extremely ill, or people who require continuous monitoring, or some of those who cannot visit the hospital or the doctor every now and then. Therefore, e-health plays a significant share, starting from detection of continuous cardiac activities (irregular or regular), sending the signal immediately to the expert, and receiving instant assistance from the doctor, so that we can avert losing lives of people. With the emergence of various wireless networks, it has become easy and convenient to transmit vital information like the ECG signals to the hospitals or to the doctors, even at remote locations and even if people are in motion or at rest. Improvements in the field of medical devices provide more accurate diagnostic instruments. Now, as the hindrances between health and health care grow fainter, new and afresh healthcare services engulf reasonable, portable, and wearable medical devices that guarantee the customers the ability to monitor their own well-being. Delivering healthcare services that are convenient and effective holds a significant position in the socioeconomic welfare of the country. Consequently, various healthcare projects with the target of providing medical benefits to the society are being pursued enthusiastically. The

B. Gandhi (✉) · N. S. Raghava
Department of Electronics and Communication Engineering, Delhi Technological University (DTU), Delhi 110042, India
e-mail: bani.gandhi@gmail.com

various techniques of measuring bio-signals include electrocardiogram (ECG) which is a dominant, standard, and powerful method used to assess cardiovascular health due to the fact that various heart conditions are reflected as disorders or abnormalities in the ECG signals. This paper discusses smart health care, specifically focusing on ECG monitoring, recording, transmitting the signals, and receiving the feedback from the doctor.

Keywords CVDs · IoT · ECG · Sensors

1 Introduction

Cardiovascular diseases have become a major cause of concern and mortality in India. The various types of CVDs like coronary heart disease that includes heart attack and cerebrovascular disease such as stroke, congenital heart disease, and rheumatic heart disease are caused due to disorders in heart and blood vessels [1]. Nearly 32% of the deaths (in 2013) globally were caused due to stroke, ischemic heart disease, and chronic obstructive pulmonary disease. Also, in India the above three mentioned heart diseases were a reason for about 32% of overall deaths [2]. If we consider the statistics for the death toll due to CVDs globally, then an estimate of 17.5 million people lost their lives in 2012. In 2015, there were about 400 million people suffering from CVDs and approximately 18 million deaths globally [3]. In India, out of 30 million heart patients, the division is such that 14 million reside in urban areas and 16 million in rural areas [4]. In a survey, it was reported that in 2015, 400 million people were suffering from various kinds of CVDs and 18 million people died because of this. It was estimated that in 2015, 7.3 million heart attacks occurred, and 110.6 million people were suffering from heart artery disease [5]. According to another report that came out in 2015, it was reported that about 31.6% people died of issues in the circulatory system [6]. The current situation of people suffering from heart diseases in India is alarming and upsetting as 'India is currently witnessing nearly two million heart attacks a year and majority of the victims are youngsters', said Dr. Ashwani Mehta, Senior Consultant Cardiologist at Sir Ganga Ram Hospital [7]. In India, about 17.5 million people die each year from CVDs. Majorly deaths are due to heart attacks and strokes, also 74% people in urban areas are at a risk of having CVDs. So roughly about 40 million people in India suffer from a CVD, out of which 19 million are from rural areas and 21 million from urban areas. Therefore, we require a health system that helps us reduce such a large number of deaths and heart patients [8]. The goal is to reduce premature CVDs globally 25% by 2025. The death toll due to CVDs has reduced, but loss of lives due to premature CVDs is still increasing [9]. All of the above presented facts provide us the conclusion that the death toll is immense due to various heart-related issues, so a healthcare system is pervasive and offers a flexible and resilient solution, which permits real-time remote health monitoring in addition with providing feedback from the doctor is required [10].

In order to get a working system where the patient's health status is recorded, it is sent to the doctor and then the feedback is provided to the caregiver, so that an immediate remedy can be provided to the patient; we need a communication system in which devices can talk to each other, so here the concept of Internet of Things (IoT) plays an important part. IoT-based healthcare systems have various advantages over conventional healthcare systems. In an IoT system, various physical devices (sensor based) are connected to each other and are linked to the Internet through various communication protocols (Bluetooth, 3G/4G, ZigBee, Wi-Fi, 6LoWPAN, etc.), so basically IoT is a network of networks or a cyber-physical network [11, 12]. The whole process of IoT is possible due to advancements in the field of Information and Communications Technologies (ICT). So a connection of information, devices, people, and processes builds up an opportunity for a system that is preventive, pervasive, affordable, and proactive. This platform enables smart devices to deliver important information without visiting the hospital or the doctor [13, 14]. Before moving on to the working system, we briefly review the concept of IoT here.

1.1 Internet of Things

As a much known fact, we all know that Internet has reformed today's communication structure and system. It has made many things convenient and affordable. This is so because the conception behind IoT is devices communicating with each other with very less or no human interaction. The IoT system basically revolves around various technologies like sensor technology, nanotechnology, embedded system technology, radio frequency identification (RFID), wireless sensor networks, etc. These are important pillars of IoT in the application of people-centric sensing [15, 16]. The ability of Internet to connect remote devices (mobile) to other devices via wireless communication protocols has bought a revolution in the domain of 'Internet', so now Internet has taken a stroll from a network of computers to a network of devices or objects, basically making the devices smart. The communication is carried out with the help of sensors introduced in the devices and various wireless technologies. It has been estimated that due to low cost, there could be about 25 billion units by 2020 and would increase by 2.1 trillion by 2025 [11, 17]. IoT has enormous applications, and as a consequence, it is making lives of people comfortable and facile [16]. As depicted in Fig. 1, the basic elements of IoT include

- Sensors: Sensors measure and collect vital parameters for example ECG, EEG, etc.
- Microprocessor/microcontroller: Used for analyzing, processing, and communication of collected data.
- Communication protocols: These protocols enable communication between various devices.

Fig. 1 IoT element
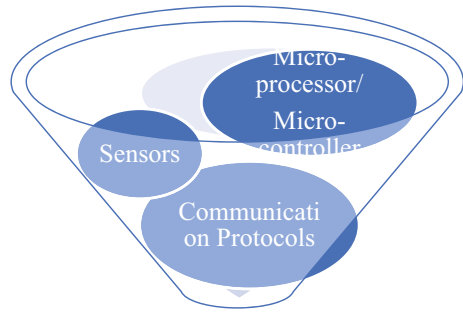
With the help of these elements, we can build an IoT system. There are various applications of IoT like smart energy management, smart security and surveillance system, smart environmental system, smart healthcare systems, and the list goes on. All thanks to the elevation in the ICT domain that has uplifted the healthcare systems. These systems ensure immediate help to the patient so that premature deaths can come to a halt and health quality can be increased. This paper discusses smart healthcare system focusing on smart ECG monitoring, recording, transmission, and reception of information, i.e., from patient to the doctor and vice versa.

IoT in Healthcare Domain. Improvisation in the healthcare domain and social well-being is the first and foremost goal of any country. IoT enables use of electronic devices or rather sensors to monitor, capture, and record data, which are then sent to the cloud or another electronic device for further usage. As a well-known fact, IoT-based healthcare systems have an edge over traditional healthcare systems. Sensor technology, which is an utmost important and fundamental principle in defining an IoT-based system, has really helped in the upliftment of this domain due to the fact that the size is decreasing, i.e., many sensors on a single chip or sensors embedded in the clothes or chairs or bed, so on and so forth. Then further, walking hands in hand with the desired wireless technologies, it brings information from the physical world to the digital world. It is expected that by 2020, IoT in the healthcare domain will be as big as $117 billion [18–20].

If we talk about the conventional way of doctor handling the patient, the patient has to visit the doctor in the hospital or the clinic. If it is a critical case, the patient might not be able to reach the destination on time; hence, the worst-case scenario is that the patient might even lose his/her life. The second issue that most of the people face is regarding the finances. In today's world, the expenses are reaching heights, so it is not always possible to visit such places every now and then just for a short visit or a consultation. The third obstacle is that continuous and long-term monitoring is not possible in conventional form of health monitoring system. Also, the availability of the doctor can vary, so considering all these issues, an IoT-enabled health monitoring system is the need of the hour so that many innocent lives can be saved, because an immediate remedy can be suggested by the doctor or the ambulance can reach the patient's place. So, an IoT supported patient can have 24 × 7 parameter(s) monitoring, and the health status of the patient can be measured by the
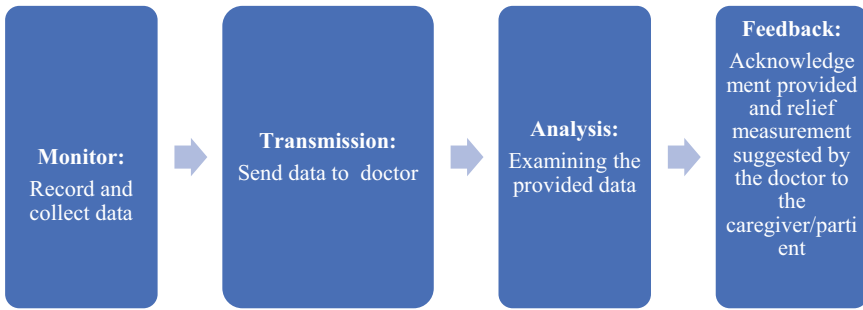
Fig. 2 Flow diagram of the process flow for IoT-based healthcare system

doctor anywhere and anytime. If the patient has access to Internet and other devices, the data generated can be stored over the cloud and can be conveniently accessed by the doctor or the caregiver [21, 22]. The abovementioned steps are explained in Fig. 2, i.e., first the data is collected, recorded, transmitted (to the doctor, or ambulance or hospital), analysis by the doctor, then the suggested remedy is sent to the patient or the caregiver.

If the above process flow is brought to a working stage, then the number of deaths due to CVDs will reduce. This is so because an instant action can be taken to retrieve the life of the affected person. A lot of people require long-term and continuous monitoring (ECG) facility, for example, expectant ladies, infants, aged/disabled people, etc. These people require remote monitoring because they cannot visit the doctor or the hospital very often. As a result, requirement of such a system becomes essential. Thus, the objectives of IoT-based healthcare systems are discussed in the next section.

Objectives of IoT in the Domain of Health Care. The main aim of IoT in the domain of healthcare is to develop, examine, authenticate, and commercialize an integrated healthcare system which is proactive and provides information (medical) from patient to the doctor and vice versa. Also, to cater to remote, continuous, long-term, and real-time monitoring along with multi-parameter (ECG, EEG, EOG, blood pressure, glucose level, etc.) detection of diseases, IoT is highly beneficial. This can further gel with the concept of self-care guidance, in which the patient with the help of smart devices can manage his own health indoor, outdoor, or at any remote location. The information can be sent to the cloud where the data can be stored and the previous reports are saved for consultation.

- Due to availability of Internet in abundance, at a lower price and most importantly, a user-friendly interface makes the system in demand in the market.
- As the world is now moving to the world of wearable(s), IoT helps in acquiring real-time information (body parameters) through intelligent sensors.
- To provide ambient intelligence.
- Analysis and prediction of premature chronic disorders through data mining techniques [23].

- One of the most important goals of IoT is more extensive interconnection of smart devices, intelligent, or non-intelligent devices [11].

Since IoT is providing us with so many applications and benefits, it does bring along various hindrances. These various challenges are discussed in the next segment.

Challenges of IoT in the Domain of Health Care. There are many obstacles that need to be eliminated before we get the desired results from the abovementioned objectives. Several challenges that exist in the arena of IoT are reviewed in this section.

- *Privacy*. Since many smart devices are involved in the process, a lot of confidential information about the patient and his health condition is either on the smart device or the cloud. So, if appropriate authentication and encryption techniques and protocols are not implemented, then various misconfigurations can take place, leading to various data protection issues, which have a huge impact on personal privacy. If the data is not protected, the consequence to this can be cross-linked data, i.e., one person can access other patient's data and this information can be misleading and may be misused. So, to avoid these situations, the data should be encrypted both at the transmission end and at the storages end.
- *Communication Protocols*. The concept of IoT relates to many heterogeneous devices connected to each other and also communicating with each other. Each type of device requires a different protocol for communicating with the other device. For instance, if we consider a scenario where ECG or EEG of a patient is being measured, the information (signals) is sent by the sensor to the device, which requires a different protocol than the one where the information (signals) is to be sent from one device to another.
- *Big Data*. Another issue that can crop up with IoT is big data in the field of health care. If a person is monitored continuously, i.e., 24×7, then large sets of data are created. All of this data may or may not be useful, thus leading to exhaustion in the bandwidth and storage spaces like cloud or the device memory. The case does not terminate here. More scenarios are there, for example, there could be various people being monitored continuously such as old/disabled people, expectant ladies, and infants and with various sensors attached to them like ECG, EEG, EOG, etc. Therefore, one should understand and gain knowledge of the relevant information and execute algorithms to discard the irrelevant data [24].
- *Energy*. Since the requirement of IoT in healthcare domain is remote, real-time, and continuous monitoring, it thus demands lots of energy to fulfill these requisites. Therefore, energy-efficient devices and algorithms are needed so that a good battery backup is maintained, and the sensor nodes are activated for a long time [25].
- *Data leakage*. This is another issue which can create a problem in the healthcare domain of IoT. Basically, the health status has to travel from one network to another or rather one device to another; this data is extremely important to the patient as well as to the doctor; hence, it is on high priority that data leakage

should be prevented, so the vendors or companies must include security into the IoT environment [24].

- *Device Vulnerability*. This issue deals with the updation and upgradation of medical devices. As a lot of devices are connected to each other, this is a complicated task due to the fact that many softwares involved in the task have version and legal issues. The problem has occurred where the network connectivity has taken a stroll from static to dynamic connectivity [22, 25].
- *Security*. Potential security implications can arise due to the fact that devices are connected to each other and to the network, so there is a great chance of the devices or the systems getting hacked. Therefore, a secure and authentic system should be in place.

The above-stated challenges need to be addressed in an appropriate manner so that we can extract major benefits from the technology. As already stated, the paper focuses on smart ECG monitoring system (IoT based); therefore, the next section discusses ECG technique to measure bio-signals.

1.2 Electrocardiogram (ECG)

Physiological signals are considered to be the first-hand information of a human being. These signals are best for capturing body signals and can very well notify about health status of an individual. To obtain the heart's signal, the most utilized technique is ECG. ECG records electrical activity of the heart, by placing electrodes on the skin at various positions like the limbs, hands, chest, etc.; and if any irregularity in the heart signal (ECG waveform) is identified, then it is an indication of one of many heart diseases [26]. The electrocardiographic supervision permits for noninvasive and persistent diagnosis, recording, and documentation of heart activities, and it is also one of the most commonly used monitoring methods [26, 27]. The ECG records the deviation of bioelectric potential with respect to time as the human heartbeats. Knowledge of analyzing and interpreting ECG signals is required for evaluating the discrepancies in the heartbeat.

In the field of medicine, ECG signals are considered to be the most useful and informative, due to the fact that they are helpful in diagnosing the cardiac-related diseases. The ECG waveform is acquired by placing the electrodes as already discussed. The ECG gives two kinds of information, one is information about the duration of the electrical wave, which decides whether the heart is functioning normally or abnormally. Second is the information on the amount of electrical activity that passes through the heart muscle, that enables to find about the parts of the heart, if they are too large [28]. The ECG waveform is shown in Fig. 3. The various deflections have been represented as P, QRS, and T. The duration/interval and amplitude of these deflections contain beneficial details regarding the nature of the heart disease [28, 29]. The electrical activity is due to depolarization and repolarization of heart
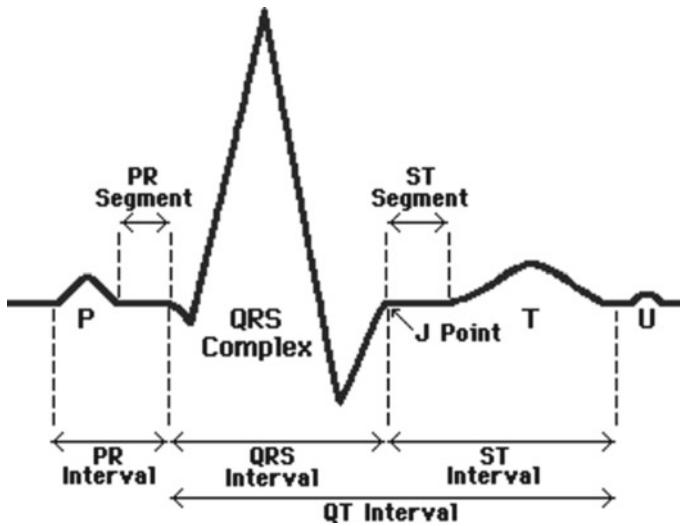
Fig. 3 ECG waveform and its deflections [30]

muscles [30]. The polarity of the ECG signal completely depends on the electrode position [28, 29].

Also the explanation of the waves, intervals, and complexes is mentioned in Table 1.

The various components of the ECG wave have been described along with their range. So, any irregularity in the abovementioned values of the wave intervals or complexes indicates that the person is ill and suffering from one of the heart diseases, for example, congenital heart disease, enlarged heart (cardiomegaly), heart muscle disease (cardiomyopathy), etc. Therefore, to predict early problems in patients, the next section illustrates the system architecture beginning with ECG signals monitoring and capturing till the patient/caregiver receives feedback from the doctor.

2 System Architecture

In this section of the paper, a smart ECG monitoring system based on IoT is presented. As depicted in Fig. 4, the preliminary step is to attain the ECG signal via sensors and then transmits ECG signals from electrodes/sensor to the smart device. The next phase is to transmit information to the medical expert's smart device, and the procedure comes at halt when the feedback is received by the patient or caregiver by the doctor. In this section, we present various modules required for the smart ECG monitoring system.

Table 1 ECG intervals and complexes

S. No.	Type	Description	Range/healthy wave (ECG interpretation)
Waves and complexes			
1	P wave	It represents atrial depolarization (left and right)	Amplitude ≤ 3 mm; upright (positive) and uniform
2	T wave	It indicates ventricular repolarization (left)	The normal T wave is usually in the same direction as the QRS except in the right precordial leads; duration < 0.2 s
3	Q wave	It is the first downward deflection after the P wave	Duration < 0.04 s; amplitude < 2.5 mm
4	R wave	It is the initial positive deflection	It has an amplitude lower than S wave
5	S wave	It is the negative deflection following the R wave	In the normal ECG, there is a large S wave in V1 (precordial or chest leads) that progressively becomes smaller, to the point that almost no S wave is present in V6 (precordial or chest leads)
6	QRS complex	It represents the ventricular depolarization (left and right)	The normal QRS complex is very variable in the frontal leads (augmented unipolar limb leads) and quite uniform in the horizontal leads (unipolar chest leads); duration: 0.06–0.12 ms
Intervals and segments			
1	PR interval	This interval starts from the P wave to the start of the QRS complex	Duration: 0.12–0.14 s
2	PR segment	It leads from end of the P wave to start of the QRS complex	A normal PR segment is present in the ECG waveform, but is invisible in case of any atrial injury: duration 50–120 ms
3	J joint	The junction between the QRS complex and ST segment	Elevation or depression of the J point is seen with the various causes of ST segment abnormality
4	QT interval	It starts from the QRS complex to the end of the T wave	Duration: 300–430 ms (males < 0.40 s; females < 0.44 s)
5	QRS interval	From the start till the end of the QRS complex	Duration ≤ 0.12 s
6	ST segment	From the end of QRS complex to the start of T wave	Duration: 80–120 ms
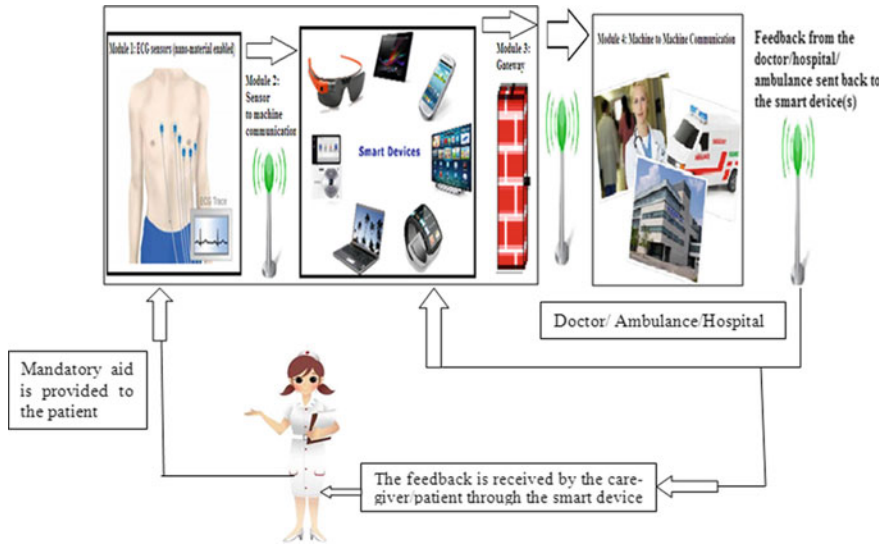
Fig. 4 ECG monitoring system based on IoT

2.1 Module 1: ECG Sensors

The foremost module of the architecture is the ECG sensors. A sensor is a device that takes input from the surroundings and then responds to it. The output then received is a signal that can be easily comprehended by a human being. Most of the sensors are readily available; sensors can be available in the raw form or embedded in the wearables. Nowadays, the sensors are embedded into the wearable(s) like vest or t-shirt, beds, wheelchairs, etc., for recording physiological signals. These sensor-based devices facilitate patients to self-monitor and self-manage disease-related issues. The foremost advantage of these devices is the elimination of the visit to the hospitals for getting the tests done which is time consuming. Also, remote and continuous monitoring is not possible and also expensive for some sections of the society. These devices are drawing attention of many researchers, since they hold a special place in the advancement in the domain of personal health management [13, 31]. Next, we consider the ECG electrodes. These are broadly classified into wet and dry electrodes.

Wet Electrodes. The most commonly used ECG electrode is the Ag/AgCl electrode that can be found in both reusable and disposable form. Ag/AgCl is referred to as the wet electrode as it requires conductive gel to maintain a good quality level of electrical conductance with the skin, so that a good quality ECG signal is received. Also, the gel has to be applied after some time to overcome signal degradation, and practically, it is not always feasible to apply the gel every now and then. Moreover, studies have shown that repetitive use of gel on the skin can cause irritation as well as dermal infection. Another drawback of using wet electrodes is that the adhesive

used in these electrodes can cause pain due to peeling of the electrode [32, 33]. In a general scenario, when the ECG has to be realized, the professional has to replace the electrodes after some hours, to avoid any kind of dermal infection or irritation. So, to avoid these issues, dry electrodes were introduced [33].

Dry Electrodes. As the name suggests, dry electrodes do not use conductive gel before the ECG monitoring process. Basically, these electrodes are designed to work without any electrolyte gel. As an alternative, dry electrodes use skin moisture or sweat. The conventional wet electrodes use the gel to penetrate through the hair to the skin to provide a conductive path, also it fills the gaps between the sensor and the skin [34]. So, dry electrodes have overcome many disadvantages of the wet electrodes.

From Table 2, we can easily infer that using dry electrodes is preferable over wet electrodes. The different types of dry electrodes are discussed below.

- *Dry Surface Electrodes*. Due to the absence of gel in the dry electrodes, the coupling between the skin and the electrode is capacitive in nature. Therefore, these electrodes cannot adjust to the skin surface that gives rise to bubbles in between, which acts as a dielectric layer. The skin–electrode impedance for dry surface electrode is more than wet electrode due to the fact that the skin and electrode contact is poor. The skin–electrode impedance is dependent on various

Table 2 Wet electrodes versus dry electrodes

S. No.	Specification	Wet electrode	Dry electrode
1	Use of gel	Yes, makes it inconvenient due to the fact that it dries with course of time	No inconvenience because no use of electrolytic gel
2	Toxicological concerns due to the gel	Yes	No
3	Motion artifacts	Yes	No
4	Initial skin preparation	Clinicians have to apply gel before starting with the ECG monitoring process	No need for such engagements
5	Continuous monitoring	Not possible because after a few hours, the gel has to be applied again because it dries up, the whole apparatus is huge, and people cannot carry it with themselves everywhere	It is possible, and dry sensors can be embedded into the wearables
6	Long-term monitoring	Not possible due to the fact already mentioned that the gel applied dries up during the course of time	Definitely possible
7	Signal degradation	Signal degrades during the course of time	No signal degradation
8	Remote monitoring	Not feasible	Possible

factors like humidity/moisture of the skin, pressure, etc. To decrease the contact impedance, we can increase the applied pressure. Dry surface electrodes are made of rigid materials that lead to high skin–electrode impedance and also high motion artifacts. Many flexible electrodes were fabricated as reported in [35–37].

- *Dry Penetrating Electrodes.* As the name suggests, these types of dry electrodes penetrate through the skin surface [38]. This method is considered to be pain free and almost noninvasive as mentioned in [39]. The needles/spikes are in direct contact with the skin since it penetrates through the dead cells of the skin [40].

- *Dry Capacitive Electrodes.* This type of sensor is a non-contact sensor that can sense the physiological signals with a gap between the body and the sensor [41, 42]. In this type of sensor, the skin and the electrode are not in direct contact with each other, they are separated either by air or by the textile, so they hold the advantage of no skin irritation and infection and are electrically safe. Examples include sensors mounted on the bed [43, 44], chairs [45], toilet seats [46], wheel chairs [47, 48], on the driving seat [49, 50], etc.

- *Nanomaterial-Enabled Dry Electrodes.* Nanotechnology is now a part of applications including medicine, electronics, food, etc., and the list is endless. In the field of electronics and computers, it has made many advancements leading to smaller, faster, and portable systems. Flexible and stretchable electronics are reaching various sectors and are being integrated into various products like wearables, aerospace, Internet of Things, etc. Nanomaterials like graphene, carbon nanotubes (CNTs) etc., are being embedded into wearables, tattoo sensors, etc., converting things into smart things. Due to various advantages of nano-materials including less motion artifacts, better wearability, and good contact with the skin, nano-material-enabled electrodes are extremely great candidates for ECG monitoring. The various types of nano-material-enabled electrodes are discussed as follows:

- *Metallic Nanomaterials (Nanoparticles/Nano-wires).* Nanoparticles/nano-wires due to their high electrical conductivity are desirable candidates for dry ECG electrodes. Nanoparticle is the most basic component, and its size varies from 10 to 100 nm. There are various types of nanoparticles like silver nanoparticles, gold nanoparticles, platinum nanoparticles, etc., and they all have different chemical and physical properties [51]. The diameter of the nano-wire is negligible, but their length can be flexible (microwires). Due to the fact that they have high ratio of surface area to volume, hence it is making them really good detectors. Nano-wires are based on a flat bed (long flat area or structure) of semiconductors like silicon or germanium. There are various types of nano-wires such as metallic nano-wire, insulating nano-wire, and semiconducting nano-wire [52]. In [37], authors have reported flexible electrodes for ECG and EEG monitoring using nanoparticles. The electrodes were made by dispersing nanoparticles in polysiloxane. It was integrated in two ways: One was in the form of flexible belt with three electrodes, and another one is a sport t-shirt with stud connectors and four electrodes. The electrodes were tested to be machine washable up to 60 °C, have proved their ability for long-term monitoring, and were comfortable. These electrodes can also be disinfected without any losses. It was observed that the ECG signals were comparable to disposable Ag/AgCl electrodes. A complete investigation was done according

to EN ISO 10993 for biocompatibility of electrodes to use on skin. The electrodes can be made with various sizes. In [53], authors have presented electrodes made of silver nano-wires (AgNWs) for ECG and EMG. These electrodes did not show any skin irritation. Silver is used in biomedical devices because of anti-bacterial properties. The AgNWs were studded below the surface of an elastomeric substrate made of PDMS (polydimethylsiloxane) that blocks the nano-wires from delimitation while fabricating a extremely high conductive surface. The electrode was stretchable and flexible, basically it conformed to the surface of the skin, thus reducing motion artifacts and reduction in the skin–electrode impedance. In some cases, AgNW electrodes performed way better than Ag/AgCl electrodes.

- *Carbon-Based Nano-materials.* There are many types of carbon-based nanomaterials like graphene, nanodiamonds, carbon nanotubes, etc. Carbon-based nanomaterials possess various properties such as high-quality thermal and electrical conductivity, high mechanical strength and also have good optical properties. Specifically talking about CNTs they possess properties such as high mechanical strength, high electrical conductivity, and most importantly can be produced in bulk at a relatively cost. In the authors have mixed CNT with PDMS; electrodes were fabricated using replica technology, and the results were comparable to Ag/AgCl electrodes. Another way of integrating CNTs with fabric is placing CNTs {Multiwall Carbon Nanotube (MWCNT)} onto the cotton fabric with tapioca starch as adhesive as mentioned in [40]. The experimental results reported proved to be better than the conventional Ag/AgCl electrodes. Also, these are low in cost, do not use hydrogel, and do not cause skin irritation.
- *Contact Penetrating Electrode.* In [54, 55], MWCNT-based electrodes were constructed in a brush-like structure, which penetrates through the outer skin layer to a depth of 10–15 μm. The results reported showed that these electrodes were robust, and a reduction in noise was observed. Also, no pain or special sensations were reported. This electrode is a brush-like structure, either coated or uncoated with silver, penetrates through the skin, and has increased contact area of the CNT.
- *Non-contact Capacitive Electrode.* It is easier to integrate electrodes with textiles. Not much work has been reported in the field of non-contact capacitive electrodes. In evaluation of screen-printed CNT-acrcylic nano-composite ECG electrodes was done. A parallel plate capacitor was formed with the CNT-acrcylic nano-composite as one of the electrodes. Screen printing technique was used for fabrication. The advantages included negligible toxic effects, flexible and stretchable, easy integration, and bulk production was possible and also easily washable.

2.2 Module 2: Sensor-to-Machine Communication

Once the ECG signals have been acquired, they have to be further sent for processing to other devices. So the next module is sensor-to-machine communication, which is discussed in this segment of the paper. The communication has to take place over a

network. It could be wired or wireless, but most of them have to happen wirelessly because of the fact that the system is IoT based. Basically, the sensor information is passed through the sensor nodes (combination of sensor, microcontroller, radio, power). The advancements in sensor technologies, networking, and signal processing have made sensors to 'smart sensors', so that they connect to various other sensors.

The networking of various sensors provides high-quality detection and measurement systems that are low cost and easy to deploy. Mostly, the sensors are connected to microcontroller for recording and transmitting the data [56]. Wireless communication eradicates the need for expensive and bulky wired networks, moving to fast and easy system deployment [54]. Here, ICT plays an important part in the communication process. ICT does not only provide an easy communication method but also a user-friendly process and also provides with long, continuous, and remote monitoring of ECG signals. Many sensors require connectivity to the sensor gateways to transmit information to the various smart devices. There are various defined protocols used for this communication. These are listed as follows:

Personal Area Network (PAN). It is a computer network around an individual person. Basically, the transmission range is less. Personal Area Network can be wired using cables like Universal Serial Bus (USB), FireWire, etc. It can also be wireless using technologies such as ZigBee, Bluetooth, RFID, so on and so forth. The network might include one or more phones, computers, printers, etc.

Local Area Network (LAN). In this network, computer networks are present at a single site for example an office building, college building, etc. A smaller area consists of two computers, and a larger may include thousands of personal computers. It can be both wireless and wired connection.

Metropolitan Area Network (MAN). This network consists of devices connected over entire city or entire campus. It is used to connect various LANs together.

Wide Area Network (WAN). It occupies a very large area, such as areas across the continents. A WAN can contain various small networks of MANs or LANs. The best example of this network is the Internet.

The data collected from the sensors has to be communicated to various other devices, so a network gateway is required for that purpose, which is discussed as the next module.

2.3   Module 3: Network Gateway

IoT is growing day by day, and billions of devices would have to be connected to each other, so a 'gateway' is required for the same, i.e., it connects networks of different architectures and environments. It repackages data according to the destination system. With the help of gateways, we are able to send the data back and forth, and the Internet would not be of any use without the gateways. Gateway is basically required to bring together networks that use different protocols. A gateway can be implemented in software or hardware or an amalgamation of both or can be virtual

[55]. It is implemented at the periphery of the network to supervise the communication process inside or outside the network. Various responsibilities of the gateways are as follows:

1. It takes heterogeneous data collected from various data sets.
2. Transformation of protocol.
3. It plays the part of both firewall and the router to provide security to the network.
4. Device connectivity.
5. Data filtering and processing.
6. Network updation and management.
7. Security for the network.
8. It acts as a bridge between the telephone network and the Internet.

A lot of data is received by the sensors, and it requires a robust system that provides a good transport system of information. Now, new IoT gateways are being delivered by various companies like Dell, Intel, Nexcom, etc. In these systems, the gateway allows to separate routine information and new information, i.e., useful information (emergency cases). These gateways are capable of performing operations and analytics on the received data [57]. The need for gateway is that the sensors do not have communication capabilities. Also, the protocols used like Bluetooth, ZigBee, 6LoWPAN, etc., used in sensor communication cannot directly communicate with larger protocols like WAN, Internet, etc., so the gateway is required for that purpose [58]. Another important thing that needs to be kept in mind is the security of the gateway which includes integrity, confidentiality, and authentication, due to the fact that gateway is attacked frequently by the hackers [59]. Now once the protocol translation has been done, the information from one smart device can be transmitted to another device. So, the next module talks about machine-to-machine communication.

2.4 Module 4: Machine-to-Machine (M2M) Communication

Since IoT is all about devices talking to each other, it is known as M2M communication. It is mandatory for machines to store large sets of data, sharing capability and have network connectivity. As the Internet has already brought a revolution, more advancements in IoT include very less or no human interaction or intervention, and the majority traffic will be generated by 'things' or 'objects'. It allows machines to communicate the that leads to self-monitoring and sending intelligent response(s) to the other device. As reported by Forbes, M2M technology is the fastest growing communication industry. The key elements in M2M communication are 'things', basically these are physical units whose characteristics, identity, and status are capable of being connected to an Internet-based system. Then a sensor is required for collection of data, and then various protocols discussed in module 2 are used for communication purpose. For wide area communication purpose, protocols such as GSM, 3G, 4G, and long-term evolution (LTE) are used. New wireless technologies for M2M communication like ultra-narrowband SIGFOX, TV white

spaces, NuelNET, etc., are emerging. M2M communication in the domain of health care plays a significant role, and it plays its role in e-health, m-health, assisted living, telemedicine, so on and so forth. For example, when a person's ECG signal is abnormal, then an alert message can be diverted to the hospital or the doctor for expert advice on the issue. This encourages in-home monitoring and makes people's lives more healthier [60]. If these systems are properly implemented and in a proper condition, it will reshape the healthcare world for the betterment in terms of accessibility, flexibility, inter-operability, and security [61]. M2M communication beholds various advantages such as low cost and effort (after the implementation) and a large massive area of communication [62].

With reference to this paper, we have divided M2M communication in two sub-categories, as discussed below:

1. **From smart device to doctor/hospital/ambulance**: The information from the sensor is sent to a smart device as already mentioned, then using one among various communication protocols, this information is sent to the other smart devices of either doctor(s), hospital(s), or ambulance(s), so that in case of emergency, necessary steps can be taken in order to help the patient.
2. **Suggestion from the doctor to the caregiver/patient**: In this particular section, the suggested remedy from the doctor is sent to the caregiver/patient so that he/she can incorporate these suggestions in chronic situations, till the doctor or the ambulance reaches the patient's destination, and this is able to save many innocent lives.

3   Conclusion

IoT is now bringing many mechanical changes in our daily lives from smart phones to smart washing machines, smart refrigerators, so on and so forth, and of course for the good. Innovations in IoT are perfect for making the world more intelligent and interactive. Same is the case with health care. With the advent of IoT in this field, it has helped the society to grow and has led to the decrease in death toll over the years. With the concept of IoT in sensors, devices, sensor networks, etc., it is creating or rather has created a smart system that can be operated anytime and anywhere. This has made the health monitoring more convenient, affordable, and rapid. Adding to these advantages, IoT-based healthcare systems also have certain features such as possibility of remote monitoring and recording. The patient is able to self-monitor and self-manage his/her own health parameters.

In this paper, the mechanism of a smart ECG system has been discussed, starting from signal acquisition (ECG), then recording of the signal (real time), sending it to the doctor, and in return getting any suggestions or remedy from the doctor. Also, for signal acquisition, nanomaterial-enabled electrodes/sensors are absolutely a fit candidate due to their various advantages over the conventional ones. This paper was only limited to monitoring of ECG signals, this can further be expanded to a platform where various other parameters can be measured like EEG, EOG, EMG,

blood pressure, glucose level, gait analysis (for fall detection), etc., so that the need for visiting the clinic or the hospital for routine checkups is minimized and the need to visit the doctor or the hospital is required only in an emergency situation. This will reduce the effort and cost of the people in the society and will lead to decrease of death toll because some diseases need immediate help and can be resolved with an instant suggestion.

References

1. World Health Organization (2017) Cardiovascular diseases. http://www.searo.who.int/india/topics/cardiovascular_diseases/en/
2. Vidya K (2014) Hypertensive heart disease top cause of deaths in India. http://www.livemint.com/Consumer/hjwtoLgWnDhlivZRUpyjiO/Hypertensive-heart-disease-top-cause-of-deaths-in-India.html
3. The Hans India (2017) One third deaths worldwide due to cardiovascular diseases. http://www.thehansindia.com/posts/index/Health/2017-05-19/One-third-deaths-worldwide-due-to-cardiovascular-diseases/301035
4. Ansari (2017) World heart day 2015: heart disease in India is a growing concern. http://food.ndtv.com/health/world-heart-day-2015-heart-disease-in-india-is-a-growing-concern-ansari-1224160
5. Hindustan Times (2017) Cardiovascular disease causes one-third of deaths worldwide: study. http://www.hindustantimes.com/lifestyle/cardiovascular-disease-causes-one-third-of-deaths-worldwide-study/story6kK2WkoTTXbYPQqCF34rFJ.html
6. Kaul R (2017) Heart diseases, stroke leading killers in India: Govt data. http://www.hindustantimes.com/india-news/heart-diseases-stroke-leadingkillers-in-india-govt-data/story-xzBCaaSM51ZpPh83adNz5I.html
7. Guirdham O (2017) Healthcare at your doorstep: USA is the largest, Japan is the oldest, and India is the fastest. http://timesofindia.indiatimes.com/life-style/health-fitness/health-news/Heart-attack-kills-one-person-every-33-seconds-in-India/articleshow/52339891.cms
8. MediConusel (2016) Current state of heart disease statistics in India. http://blog.medicounsel.com/2016/09/12/heart-disease-statistics-india-2016/
9. Roth GA, Huffman MD, Moran AE, Feigin V, Mensah GA, Naghavi M, Murray CJL (2015) Global and regional patterns in cardiovascular mortality from 1990 to 2013, vol 132, no 17. American Heart Association, Inc., pp 1667–1678
10. Tabish R, Ghaleb AM, Hussein R, Touati F, Mnaouer AB, Khriji L, Rasid MFA (2014) A 3G/WiFi-enabled 6LoWPAN-based U-healthcare system for ubiquitous real-time monitoring and data logging. In: 2014 Middle East conference on biomedical engineering (MECBME), pp 277–280
11. Darshan KR, Anandakumar KR (2015) A comprehensive review on usage of internet of things (IoT) in healthcare system. In: International conference on emerging research in electronics, computer science and technology, pp 132–136
12. Wang C, Daneshmand M (2013) Guest editorial special issue on internet of things (IoT): architecture, protocols and services. IEEE Sens J 13(10):3505–3510
13. Lake D, Milito R, Morrow M, Vargheese R (2015) A custom internet of things healthcare system. In: 10th Iberian conference on information systems and technologies—CISTI'2015, Águeda, Portugal, vol 1, pp 653–658
14. Venkatramanan P, Rathina I (2014) Healthcare leveraging internet of things to revolutionize healthcare and wellness. IT Services Business Solutions Consulting, © 2014 Tata Consultancy Services Limited

15. Sendra S, Granell E, Lloret J (2013) Smart collaborative mobile system for taking care of disabled and elderly people. Mobile Netw Appl 19(3):1–16
16. Yang L, Ge Y, Li W, Rao W, Shen W (2014) A home mobile healthcare system for wheelchair users. In: Proceedings of the 2014 IEEE 18th international conference on computer supported cooperative work in design, pp 609–614
17. Iyer R, Mishra R (2014) Building intelligent internet of things applications using microsoft stream insight. IGATE Global Solution, pp 1–7
18. McCue TJ (2015) $117 billion market for internet of things in healthcare by 2020. http://www.forbes.com/sites/tjmccue/2015/04/22/117-billionmarket-for-internet-of-things-in-healthcare-by-2020/#58d793712471
19. Kevdar J (2016) Harnessing the internet of health things. http://www.himssasiapac.org/sites/default/files/HIMSSAP_ThematicReport_HarnessingtheInternetofHealthThings.pdf
20. William PAH, McCauley V (2016) Always connected: the security challenges of the healthcare internet of things. In: 2016 IEEE 3rd world forum on internet of things (WF-IoT), pp 30–35
21. Dhar SK, Bhunia SS, Mukherjee N (2014) Interference aware scheduling of sensors in IoT enabled health-care monitoring system. In: Fourth international conference of emerging applications of information technology. IEEE, pp 152–157
22. Karam A (2015) Internet of things: objectives and scientific challenges. https://www.linkedin.com/pulse/internet-things-objectives-scientific-challenges-ahmed-karam%D8%A3%D8%AD%D9%85%D8%AF-%D9%83%D8%B1%D9%85
23. Vyas DA, Bhatt D, Jha D (2016) IoT: trends, challenges and future scope. IJCSC 7:186–197
24. Vani G, Bharathi Malakreddy A (2016) Security challenges in IoT applications in health care domain. Int J Adv Electron Comput Sci 141–144
25. Willaims PAH, Woodward AJ (2015) Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. Med Devices Evid Res 305–316
26. Sanamdikar ST, Hamde ST, Asutka VG (2015) A literature review on arrhythmia analysis of ECG signal. Int Res J Eng Technol (IRJET) 2(3):307–312
27. Wang J (2016) Potential solutions for managing real-time ECG/arrhythmia monitoring alarms: a review. In: Computing in cardiology conference (CinC)
28. Kundu M, Nasipuri M, Basu DK (2000) Segmentation of plane curves identification of sudden cardiac arrest, knowledge-based ECG interpretation: a critical review. Pattern Recogn 33:351–373
29. Horowitz SL (1975) A syntactic algorithm for peak detection in waveforms with applications to cardiography. Commun ACM 5:281–285
30. Rondoni J (2002) Paroxysmal atrial fibrillation and electrocardiogram predictor. MAS.622J
31. Aiboud Y, Mhamdi JE, Jilbab A, Sbaa H (2016) Review of ECG signal de-noising techniques. In: 2015 third world conference on complex systems (WCCS)
32. Meziane N, Webster JG, Attari M, Nimunkar AJ (2013) Dry electrodes for electrocardiography. Physiol Meas 34:R47–R69
33. Kwak MK, Jeong H-E, Suh KY (2011) Rational design and enhanced biocompatibility of a dry adhesive medical skin patch. Adv Mater 23:3949–3953
34. COGNIONICS, INC (2011–2017) Comparing cognionics dry with conventional wet sensors. http://www.cognionics.com/index.php/technology/dry-electrode-systems
35. Hoang MV, Chung HJ, Elias AL (2016) Irreversible bonding of polyimide and polydimethylsiloxane (PDMS) based on a thiol-epoxy. J Micromech Microeng 26(10):1–9
36. Meng Y, Li Z, Chen J (2016) A flexible dry electrode based on APTES-anchored PDMS substrate for portable ECG acquisition system. Microsyst Technol 22(8):2027–2034
37. Hoffmann KP, Ruff R (2007) Flexible dry surface-electrodes for ECG long-term monitoring. In: Annual international conference of the IEEE engineering in medicine and biology proceedings, vol 4, pp 5739–5742
38. Liao LD, Wang I-J, Chen SF, Chang JY, Li CT (2011) Design, fabrication and experimental validation of a novel dry-contact sensor for measuring electroencephalography signals without skin preparation. Sensors 11:5819–583

39. Hsu LS, Tung SW, Kuo CH, Yang YJ (2014) Developing barbed microtip-based electrode arrays for biopotential measurement; micromachined electrodes for biopotential measurements. Sensors 14:12370–12386
40. Yao S, Zhu Y (2016) Nanomaterial-enabled dry electrodes for electrophysiological sensing: a review. JOM 68(4):1145–1155
41. Chi YM, Cauwenberghs G (2009) Micropower non-contact EEG electrode with active common-mode noise suppression and input capacitance cancellation. In: 31st annual international conference of the IEEE EMBS, pp 4218–4221
42. Harland CJ, Clark TD, Prance RJ (2001) Electric potential probes—new directions in the remote sensing of the human body. Meas Sci Technol 13(2):163–169
43. Ishijima M (1993) Monitoring of electrocardiograms in bed without utilizing body surface electrodes. IEEE Trans Biomed Eng 40(6):593–594
44. Ueno A, Akabane Y, Kato T, Hoshino H, Kataoka S, Ishiyama Y (2007) Capacitive sensing of electrocardiographic potential through cloth from the dorsal surface of the body in a supine position: a preliminary study. IEEE Trans Biomed Eng 54(4):759–766
45. Aleksandrowicz A, Leonhardt S (2007) Wireless and non-contact ECG measurement system—the Aachen smartchair. Acta Polytech 2:68–71
46. Kim KK, Lim YK, Park KS (2004) The electrically noncontacting ECG measurement on the toilet seat using the capacitively-coupled insulated electrodes. In: Conference proceedings of IEEE engineering in medicine and biology society, vol 4, pp 2375–2378
47. Postolache O, Girao PS, Mendes J (2009) Unobstrusive heart rate and respiratory rate monitor embedded on a wheelchair. In: IEEE international workshop on, medical measurements and applications, MeMeA, pp 83–88
48. Milenković A, Milosevic M, Jovanov E (2013) Smartphones for smart wheelchairs. In: 2013 IEEE international conference on body sensor networks, BSN 2013
49. Ji Q, Zhu Z, Lan P (2004) Real-time nonintrusive monitoring and prediction of driver fatigue. IEEE Trans Veh Technol 53(4):1052–1068
50. Singh RK, Sarkar A, Anoop CS (2016) A health monitoring system using multiple non-contact ECG sensors for automotive drivers. IEEE, pp 1–6
51. Horikoshi S, Serpone N (2013) Introduction to nanoparticles, microwaves in nanoparticle synthesis. Wiley-VCH Verlag GmbH & Co. KGaA, pp 1–23
52. Aggarwal T. Nanowire—applications and advantages. Elprocus. https://www.elprocus.comf/nanowire/
53. Myers AC, Huang H, Zhu Y (2015) Wearable silver nanowire dry electrodes for electrophysiological sensing. RSC Adv 5(15):11627–11632
54. Mason A, Yazdi N, Chavan AV, Najafi K, Wise KD (1998) A generic multielement microsystem for portable wireless applications. Proc IEEE 6(8):1733–1746
55. Mitchell M (2017) Internet and network. https://www.lifewire.com/definition-of-gateway-817891
56. Zhou J, Mason A (2002) Communication buses and protocols for sensor networks. Sensors 2:244–257
57. Treadaway J (2017) Using an IoT gateway to connect the "things" to the cloud. http://internetofthingsagenda.techtarget.com/feature/Using-an-IoT-gateway-to-connect-the-Things-to-the-cloud
58. Colosimo C (2017) Reducing costs by prototyping with service virtualizing. https://dzone.com/articles/reducingcosts-by-prototyping-withservice-virtual
59. Desai N (2016) Global sign blog. https://www.globalsign.com/en/blog/what-is-an-iot-gateway-device/
60. mHealth Alliance (2017) A collection of seminal reports from the mHealth alliance supporting national digital health systems. http://www.mhealthknowledge.org/resource-type/mhealth-alliance
61. http://www.zdnet.com/article/m2m-and-the-internet-of-things-a-guide/M2M and the internet of things: a guide

62. Patel KL, Patel SM (2016) Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application and future challenges. Int J Eng Sci Comput 6(5):6122–6131
63. Liu B, Chen Y, Luo Z, Zhang W, Tu Q, Jin X (2015) A novel method of fabricating carbon nanotubes-polydimethylsiloxane composite electrodes for electrocardiography. J Biomater Sci Polym Ed 26(16):1229–1235
64. Kumar PS, Rai P, Oh S, Kwon H, Varadan VK (2012) Nanocomposite electrodes for smartphone enabled healthcare garments: E-bra and smart vest. In: Nanosystems in engineering and medicine, Proceedings of SPIE, vol 8548, pp 85481O-1–85481O-8

Ensemble Learning Techniques and Their Applications: An Overview

Anil Kumar Dasari, Saroj Kr. Biswas, Dalton Meitei Thounaojam, Debashree Devi, and Biswajit Purkayastha

Abstract Despite of important successes in knowledge discovery, conventional machine learning techniques may fail to get satisfactory achievements when dispensing with complex and noisy data, like high-dimensional, imbalanced, etc. This is because achieving many features and the underlying structure of data is difficult with these methods. Ensemble learning, on the other hand, tries to combine data fusion, modeling, and mining into a coherent framework as a single research hotspot. Specifically, ensemble model generates a set of diverse classifiers and then creates a robust prediction model by combining the diverse classifier and creates a robust and reliable prediction model. Therefore, this paper systematically discusses ensemble learning methods and analyzes them critically. This paper also reports their suitability in different applications and their classifications. The paper targets the beginners in ensemble learning to provide a precise and compact view about ensemble learning methods.

1 Introduction

In the last two decades, technology has improved significantly. Most of the research works are going on in Artificial Intelligence (AI) to reduce human intervention, where machines can take their own decision to perform any task. Machine learning

A. K. Dasari (✉) · S. Kr. Biswas · D. M. Thounaojam · B. Purkayastha
Department of Computer Science and Engineering, National Institute of Technology Silchar, Silchar, India
e-mail: anil_rs@cse.nits.ac.in

D. M. Thounaojam
e-mail: dalton@cse.nits.ac.in

D. Devi
Department of Computer Science and Engineering, Indian Institute of Information Technology Guwahati, Guwahati, India

(ML) has now become an integral part of AI. Different ML algorithms are used to solve various domains like computer vision, image processing, the Internet of Things (IoT), etc. ML algorithms follow different learning paradigms to solve a particular problem, and based on those learning paradigms, supervised learning, semi-supervised learning, unsupervised learning, and reinforcement learning are the four types of machine learning algorithms. Training using labeled data is required for supervised learning. The input value and the desired target output value are both included in each labeled training dataset. In semi-supervised learning, the training dataset includes both labeled and unlabeled data. The classifiers are trained to study the patterns to analyze and label the data as well as to predict the unknown labels. Unsupervised machine learning extracts hidden insights from unlabeled datasets, such as cluster analysis. Reinforcement learning, the fourth category, is a process for learning behavior via responses collected through communications with external conditions [1]. In terms of data processing, both unsupervised and supervised learning methods are employed for data analysis, with reinforcement learning being the favored way for decision-making tasks [2].

The aim of supervised learning methods is to find the relationship between input and target attributes which are mostly applied to straightforward tasks that involve labeled datasets. There are various supervised learning classifiers like Naïve Bayes, decision tree, support vector machine, random forest, K-nearest neighbor, etc. There are various limitations encountered for single classifier, such as when the data is unstructured and complex and has large number of features [3], when the problem is nonlinear, when the data is high-dimensional, and when the dataset size is very small which leads to over-fitting [4] and non-interpretability of the training data [5]. In case of non-interpretable training data, for employing ensemble learning model such as random forest, the number of trees has to be chosen manually.

The principal concept of ensemble learning is to evaluate multiple individual pattern classifiers and incorporate their predictions to have a classifier that outperforms each. It is vital to have a more authentic ensemble of classifiers than its classifiers to have an accurate ensemble of classifiers [6]. The error rate of an accurate classifier is lower than that of a random guess on unknown values. The failure of a single learning algorithm leads to the development of ensemble classifiers for three primary reasons: statistical, computational, and representational [6]. Because ensemble models are more stable and, more crucially, predict better than single classifiers, they have been widely used in credit scoring applications and other domains [7]. To overcome the drawbacks of a single classifier system on imbalanced or unstructured data, intelligent and self automatically updatable machine learning systems are very much required. Since machine learning is a system that incrementally learns concerning change in environment and ensemble of the classifier is a technique under machine learning that answers such challenges and addresses such issues. Ensemble learning improves machine learning performance by merging many models. In contrast to a single machine learning (ML) model, this technique provides greater predictive performance. We, as a human behavior, used to balance and combine individual ideas to reach our final decision [8]. The ensemble learning

classifier simulates this human behavior, considering many opinions before making an important decision.

2   Classification of Ensemble Learning Techniques

There are various ensemble learning techniques such as boosting, bagging, stacking, Mixture of Experts (MoE), Error-Correcting Output Codes (ECOC), and Cluster Ensembles.

2.1   Boosting

Boosting is a form of ensemble learning that turns a group of weak learners into strong ones. Moreover, this framework provides a "weak" version, the most efficient and imperceptibly above random view, to be randomly nudged into an accurate and robust appearance. There are three boosting algorithms such as AdaBoost (Adaptive Boosting) algorithm, Gradient Boosting (GB) algorithm, and eXtreme Gradient Boosting (XGB) algorithm.

AdaBoost (Adaptive Boosting) Algorithm
The most well-known and reliable technique for constructing an ensemble design is AdaBoost [9]. AdaBoost's primary purpose is to concentrate on samples that were misclassified before while using a different inducer. A weight assigned to each sample in the training set determines the level of focus delivered. Random weights are initialized to each of the samples in the first iteration, which is calculated as 1/Number of training samples. For every iteration, the weights of perfectly coordinated samples are reduced while the results of misclassified examples are increased. The unique base learners are also given weights based on their total prediction ability. The pseudo-code for ensemble approach of AdaBoost is presented below:

Algorithm 1 The Adaboost Algorithm

> **Input**: I (a weak inducer), T (the number of iterations), S (training set)
> *Output*: $M_t, \alpha_t; \forall t = 1, \dots, T$
> **for** each t in $1, \dots, T$ **do**
> Build classifier M_t using I and distribution D_t
> $\varepsilon_t \leftarrow \sum i: M_t(x_i) \neq y_i \, D_t(i)$
> *If* $\varepsilon_t > 0.5$ *then*
> $T \leftarrow t - 1$
> **exit loop**
> **end**
> **else**

$$\alpha_t \leftarrow 1/2\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$$

$$D_{t+1}(i) = (i).e^{-\alpha_t y_i M_t(x_i)}$$

Normalize D_{t+1} to be a proper distribution

t++

end

end

The AdaBoost algorithm is mostly used to research and solve classification problems. For example, literature is used to answer the two-class problem, multi-class single-label problem, multi-class multi-label problem, two-class single-label problem, and regression problem. This paper covers the basic theory and implementation of the two-class single-label issue in text categorization using AdaBoost. It looks at the AdaBoost.M1 and AdaBoost.M2 algorithms and how they are used to solve the multi-class single-label problem, as well as the AdaBoost.MR and AdaBoost.MH algorithms and how they are used to solve the multi-class multi-label problem.

AdaBoost.M1 and AdaBoost.M2 are two versions of the algorithm that we present. For binary classification issues, the two versions are identical; the only difference is in their applications.

AdaBoost.M1 and AdaBoost.M2 models are extensions to multi-class classifications (with M2 overcoming a restriction on the maximum error weights of classifiers from M1).

AdaBoost.M1

AdaBoost was initially known as AdaBoost.M1 by the technique's creators, Freund and Schapire. It has been termed discrete AdaBoost in recent years because it is applied for classification rather than regression. The performance of any machine learning technique can be improved by using AdaBoost. It is helpful for weak learners. These are models that produce accuracy that is somewhat above random chance on a classification problem. Decision trees with one level are the best-suited and hence most widely used algorithm with AdaBoost. Decision stumps get their name from the fact that these trees are so short and only have one decision for classification.

Algorithm 2 AdaBoost.M1

Input: Sequence of m examples $((x_1, y_1), \ldots, (x_m, y_m))$ with labels $y_i \in Y = \{1, \ldots, k\}$

Weak learning algorithm WeakLearn Integer T specifying number of iterations

Initialize $D_1(i) = 1/m$ for all I

Do for $t = 1, 2, \ldots T$

1. Call WeakLearn, providing it with distribution D_t
2. Get back hypothesis $h_t: X \rightarrow Y$
3. Calculate the error of h_t: $\in_t = \sum i: h_t(X_i) \neq y_i D_t(i)$. If $\in_t > 1/2$, *then set $T = t - 1$ and abort loop*
4. *Set $\beta_t = \in_t/(1 - \in_t)$*

5. *Update distribution D_t*: $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} \beta_t & \text{if } h_t(x_i) = y_i \\ 1 & \text{otherwise} \end{cases}$

where Z_t is a normalization constant (chosen so that D_{t+1} will be a distribution)

Output: the final hypothesis: $h_{\text{fin}}(x) = \underset{y \in Y}{\arg\max} \sum_{t:h_t(x)=y} \log \frac{1}{\beta_t}$

AdaBoost.M2

By increasing the connection between the boosting algorithm and the weak learner, the second version of AdaBoost seeks to address the challenge of processing weak hypotheses with error greater than 0.5. AdaBoost.M2 is a multi-class boosting technique for problems with weak base classifiers. The goal of the algorithm is to reduce the training error to a very low level. The boosting technique AdaBoost.M2, which is based on these approaches, is found to achieve boosting if each weak hypothesis has a pseudo-loss that is somewhat better than random guessing.

Algorithm 3 AdaBoost.M2

Input: Sequence of m examples $((x_1, y_1), \ldots, (x_m, y_m))$ with labels $y_i \in Y = \{1, \ldots, k\}$

Weak learning algorithm WeakLearn Integer T specifying number of iterations

Let $B = \{(i, y): i \in \{1 \ldots m\}, y \neq y_i\}$

Initialize $D_1(i, y) = 1/|B|$ for $(i, y) \in B$

Do for $t = 1, 2, \ldots T$

1. Call **WeakLearn**, providing it with mislabel distribution D_t.
2. Get back hypothesis $h_t: X \times Y \to [0, 1]$
3. Calculate the pseudo-loss of h_t: $\epsilon_t = \frac{1}{2} \sum_{(i,y) \in B} D_t(i, y).(1 - h_t(x_i, y_i) + h_t(x_i, y))$
4. *Set $\beta_t = \epsilon_t / (1 - \epsilon_t)$*
5. Update D_t: $D_{t+1}(i, y) = \frac{D_t(i, y)}{Z_t} . \beta_t^{(\frac{1}{2})(1 + h_t(x_i, y_i) - h_t(x_i, y))}$
6. where Z_t is a normalization constant (chosen so that D_{t+1} will be a distribution)
7. **Output**: the final hypothesis: $h_{\text{fin}}(x) = \underset{y \in Y}{\arg\max} \sum_{t=1}^{T} \left(\log \frac{1}{\beta_t} \right) h_t(x, y)$

AdaBoost.MH (AdaBoost with Multi-class Hamming Loss)

Schapire and Singer [10] created AdaBoost.MH. AdaBoost MH is a multi-class technique that uses the Hamming loss, which is the average exponential loss on L binary classification tasks. "Is the proper label Y or one of the other labels?" for each case x and each alternative label y and for each potential label y and each possible case x.

Algorithm 4 AdaBoost.MH

Input Data: $D = \{(x^{(i)}, Y^{(i)})\}^N$, where $x^{(i)} \in \mathfrak{X}$ and $Y^{(i)} \in 2^{\mathfrak{y}}$ for $i = 1, \ldots N$, where $\mathfrak{y} = \{1, \ldots, K\}$ is a set of K elements

Total number of rounds: T

Initialize: Define the initial probability distribution $D_1(i) = \frac{1}{NK}$ for $i = 1, \ldots, N, \ell = 1, \ldots, K$.

Do for $t = 1, 2, \ldots, T$

• Train using the probability distribution D_t on $\{1, \ldots, N\} \times \{1, \ldots K\}$.

• Get a hypothesis (classifier) $\bar{h} : \mathfrak{X} \times \mathfrak{Y} \to \{-1, 1\}$, equivalently $h : \mathfrak{X} \to 2^{\mathfrak{Y}}$

• Choose $\alpha_t \in \mathbb{R}$

• update the probability distribution

$D_{t+1}(i, \ell) = \frac{D_t(i) \exp(-\alpha_t y_\ell^i \bar{h}_t(x^i, \ell))}{Z_t}$,

where $Z_t = \sum_{t, \ell} D_t(i, \ell) \exp\left(-\alpha_t y_\ell^{(i)} \bar{h}_t(x^{(i)}, \ell)\right)$

$(Note: e_H(h_t(x^{(i)}), Y^{(i)}) = \left|\left\{\ell : y_\ell^{(i)} \bar{h}_t(x^{(i)}, \ell) = -1\right\}\right|$

Output: the final hypothesis (Classifier)

AdaBoost.MR (AdaBoost with Multi-class Ranking Loss)

Schapire and Singer [10] created AdaBoost.MR (AdaBoost with Multi-class Ranking Loss). They concluded that the performance improvements were due to the first few classifiers and that the classifier could achieve effective performance with fewer classifiers rather than creating binary problems, such as "Is the correct label *y* or *y*1?" for sample *x* with correct label *y* and each incorrect label *y*1.

Algorithm 5 AdaBoost.MR

Input Data: $D = \{(x^{(i)}, Y^{(i)})\}^N$, where $x^{(i)} \in \mathfrak{X}$ and $Y^{(i)} \in 2^{\mathfrak{y}}$ for $i = 1, \ldots N$, where $\mathfrak{y} = \{1, \ldots, K\}$ is a set of K elements usually identified by with $\{1, \ldots K\}$.
Total number of rounds: T

Initialize: Define the initial probability distribution D_1 by

$D_1(i, \ell_0, \ell_1) = \begin{cases} \frac{1}{N|\mathfrak{Y}-Y^{(i)}||Y^{(i)}|} & \text{if, } \ell_0 \notin Y^{(i)} \text{ and } \ell_1 \in Y^{(i)} \\ 0 & \text{else} \end{cases}$

Do for $t = 1, 2, \ldots T$:

• Train using D_t

• *Get a hypothesis (ranking function)* $h_t \to \mathfrak{X} \times \mathfrak{y} \to \mathbb{R}$

• Choose α_t

• update the probability distribution

$D_{t+1}(i, \ell_0, \ell_1) = \frac{D_t(i, \ell_0, \ell_1) \exp\{\frac{1}{2}\alpha_t[h_t(x^{(i)}, \ell_0) - h_t(x^{(i)}, \ell_1)]\}}{Z_t}$,

where $Z_t = \sum_{i, \ell_0, \ell_1} D_t(i, \ell_0, \ell_1) \exp\{\frac{1}{2}\alpha_t[h_t(x^{(i)}, \ell_0) - h_t(x^{(i)}, \ell_1)]\}$

Output: the final hypothesis

$f(x, \ell) = \sum_t \alpha_t h_t(x, \ell)$

Gradient Boosting (GB) Algorithm

Gradient Boosting (GBM), developed by Friedman [11], is a method of combining predictors sequentially into an ensemble, with past predictors assisting their successors by enhancing the representation's accuracy. New predictors mitigate the effects

of previous predictor's errors. The gradient boosted maintains the gradient in recognizing and correcting errors in learner's predictions.

Algorithm 6 Gradient Boosting Algorithm

Input: training set $\{(x_i, y_i)\}^n$, a differentiable loss function $L(y, F(x))$, number of iterations M

1. Initialize model with a constant value

$F_0(x) = \underset{\gamma}{\arg\min} \sum_{i=1}^{n} L(y_i, \gamma).$

2. For $m = 1$ to M

• Compute so-called pseudo-residuals:

$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}$ for $i = 1, \ldots$

Fit a base learner (or weak learner, e.g. tree) closed under scaling $h_m(x)$ to pseudo residuals, i.e. train it using the training set $\{(x_i, r_{im})\}_{i-1}^{n}$

• Compute the multiplier γ_m by solving the following one dimensional optimization problem:

$\gamma_m = \underset{\gamma}{\arg\min} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)).$

Update the model

$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$

Output $F_M(x)$

XG Boosting (Extreme Gradient)

XGBoost [12] is a scalable machine learning method for tree boosting that has increased demand in recent years among machine learning practitioners. XGBoost provides the use of decision trees with boosted gradients, implementing enhanced speed and execution. It relies massively on the computational speed and the appearance of the target representation. Model training should follow a sequence, thus making the implementing gradient boosted machines slow.

Algorithm 7 XG Boosting

Input: **Dataset D** $= \left\{ (x_i, y_i) : i = 1 \ldots n, \ x_i \in \mathbb{R}^m, \ y_i \in \mathbb{R} \right\}$

we have n samples with m features

1. The prediction value model is $\hat{y}_i \sum_{k=1}^{K} f_k(x_i), \ f_k \in F$

where f_k is independent regression tree and $f_k(x_i)$ is prediction score given by kth tree to ith sample

2. The set of functions f_k in the regression tree model can be learned by minimizing the objective function:

Obj $= \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$ where l is training loss function

3. To avoid over-fitting, the term Ω penalizes the complexity of the model:

$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda ||w||^2$

where γ and λ are the degrees of regularization. T and w are the numbers of leaves

4. *Let $\hat{y}_i^{(t)}$* be the prediction of the *i*th instance at the *t*th iteration it needs to add f_t to minimize the following objective:

$$\text{Obj}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t)$$

5. To remove the constant term following equation given

$$\text{Obj}^{(t)} = \sum_{i=1}^{n}\left[g_i f_t(x_i) + \tfrac{1}{2}h_i f_t(x_i)^2\right] + \Omega(f_t)$$

where $g_i = \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ *and* $h_i = \partial^2_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ are the first *and the second order gradient on l*

6. Then the objective is rewritten as:

$$\text{Obj}^{(t)} = \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2}h_i f_t(x_i)^2\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2$$

$$= \sum_{j=1}^{T}\left[\left(\sum_{i\in I_j} g_i\right)w_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)w_j^2\right] + \gamma T$$

where $I_j = \{i|q(x_i) = j\}$ denotes the instance set of leaf j, For a fixed tree structure the optimal weight w_j^* of leaf j

7. The corresponding optimal value can be calculated by

$w_j^* = -\frac{G_j}{H_j+\lambda}$

$\text{Obj}^* = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j+\lambda} + \lambda T$

where $G_j = \sum_{i\in I_j} g_i$, $H_j = \sum_{i\in I_j} h_i$, obj presents the quality of a tree *structure q*.

2.2 Bagging

Bagging (bootstrap aggregation [13]) is a simplistic yet effective method for constructing a set of independent models. As a replacement, each classifier is trained using a subset of the original dataset's cases. To provide a sufficient number of training examples, each sample usually comprises the same number of cases as the original dataset. The final prediction of an unseen model is determined by the majority voting of the classifiers [14]. Because replacement sampling is used, some of the original examples are likely to appear more than once when training the same classifier, while others may be excluded entirely. Bagging can be easily accomplished in parallel by training each inducer using different processing units because the inducers are presented independently. There are two types of bagging techniques random forest and bagging meta-estimator algorithm.

Random Forest

Random forest [15] is a method for increasing decision tree classifier variation by training each classifier on the entire dataset in a rotating feature space The characteristics are randomly separated into K subsets, and Principal Component Analysis

(PCA) is employed for each subset while constructing the base classifier's training data. The aim behind PCA is to convert all possible correlated data into a set of linearly uncorrelated features (called principal components). Each part is a linear combination of the previous one. The modification ensures that the first principal component has the greatest amount of variation feasible. Under the restriction of being orthogonal to the previous components, each following element has the largest variance achievable.

Algorithm 8 Random Forest Algorithm

> **Input**: IDT (a decision tree inducer), T (the number of iterations), S (training set), μ (*the subsample size*), N (*number of attributes used in each node*)
> **Output**: M_t: $\forall_t = 1, \dots, T$
> **for** *each t in* 1, ..., *T do*
> $S_P \leftarrow$ Sample μ instances from S with replacement Build classifier M_t
> using IDT (N) *on S_t*
> t++
> End

Bagging Meta-estimator

Bagging methods are a type of ensemble algorithm that uses numerous instances of a black-box estimator on random subsets of the initial training set and then combines their guesses to get a final prediction. These methods are used to lower the variance of a base estimator (e.g., a decision tree) by introducing randomization into the construction mechanism and then constructing an ensemble from it. In many circumstances, bagging approaches provide a simple way to improve a single model without having to change the underlying base algorithm.

Random subsets are created from the original dataset (bootstrapping).

Algorithm 9

Step 1: Let N be the size of the training set.

Step 2: for each of t iterations:

Step 3: sample N instances with replacement from the original training set

Step 4: apply the learning algorithm to the sample

Step 5: store the resulting classifier

Step 6: for each of the t classifiers:

Step 7: predict class of instance using classifier

Step 8: return class that was predicted most often.

2.3 Stacking

Using a meta-classifier or meta-model, this strategy also integrates various classifications or regression algorithms. Lower-level models are trained using the entire training dataset, and the composite model is then trained using the results of the lower-level models. In contrast to boosting, each lower-level model is trained in parallel. The training dataset for the next model is the prediction from the lower-level models, forming a stack in which the top layer of the model is more trained than the bottom layer of the model. The top layer model is based on lower-level models and has a high prediction accuracy. The stack grows until the best prediction is made with the least amount of inaccuracy.

The prediction of the combined model or meta-model is based on the predictions of the many weak models or lower layer models. It focuses on producing a less biased model. Stacking Generalization is a technique used in ensemble learning. The meta-model is built from bootstrap samples using a dataset, and the outputs of the exclusion dataset are used as inputs. The most fundamental fashions are Base Model Levels 0 and 1. The level 1 model's goal is to integrate the output set in order to correctly identify the target, correcting the level 0 models' errors.

2.4 Mixture of Experts (MoE)

One of the most popular and interesting combining approaches for improving machine learning performance is the Mixture of Experts (MoE) [16]. The ME is built on the divide-and-conquer approach, in which the problem space is divided among a few neural network specialists monitored by a gating network. Several techniques were developed to split the problem space across the specialists in previous research on ME. We categorize the ME literature based on this distinction to better examine and assess these strategies. ME implementations are divided into two groups: The first is a mix of implicitly localized experts, and the second is a mix of implicitly localized experts (MILE).

Mixture of Implicitly Localized Experts (MILE)
Through a competitive learning strategy, many distinct experts learn to manage several but overlapped subsets of training data in the typical ME outlined by Jacobs et al. [17]. For each training scenario, a gating network chooses which experts should practice. It compensates the expert(s) with the most dependable performance with more effective mistake feedback indicators. The gating network divides the input space into subspaces based on the experts' knowledge and identifies a learning subspace for each expert.

Mixture of Explicitly Localized Experts (MELE)
Since 2000, researchers have proposed MELE rules, which often have more representation than MILE approaches [18, 19]. Despite having the same components as

the MILE techniques, such as some specialists and a gating network, these systems act independently as partitioning devices. MELE approaches partition the input space into more divisible spaces than MILE, which stochastically partitions each expert network's input space and practices on layered and stochastic input space areas. Each expert is then focused on a pre-determined subspace, with updated ME learning rules.

2.5 Error-Correcting Output Codes (ECOC)

Error-Correcting Output Codes is an ensemble learning technique. It applies to an issue in more than one training and divides it into several binary problems. If there are T models in the set, every class is encoded first as a two-string length T. After that, each version attempts to distinguish a subset of the original instructions from the rest. For example, we might examine the model to distinguish between "A Wonderful A" and "No Longer A Great." With T models, we get a binary series of T duration after the predictions. The elegance that denotes this is closest to this binary series is the group's final decision.

2.6 Cluster Ensembles

A clustering ensemble intends to connect multiple clustering models to provide a more reliable result than the unique clustering algorithms in terms of flexibility and quality. Cluster Ensembles are a method of ensemble learning that is unsupervised. The idea is to produce numerous alternative clustering's of a dataset, maybe using various algorithms, and then combine the opinions of other groups into a single result. In theory, the final ensemble clustering model might be more reliable than using a single classifier.

3 Different Kinds of Applications of Ensemble Learning

In this phase, we examine the four types of ensemble learning applications that have dominated the field during the last decade. They are face and fingerprint identification, remote sensing, one versus all recognition (which covers a wide range of fault detection and intrusion), and medicine. Every one of these professions has its own set of analytical challenges (from an excessive number of records "to too few stats of a particular type"), making it difficult to implement standard pattern popular algorithms and thus are fertile ground for ensembles' applications.

There are various types of applications for ensemble learning methods which are listed below.

1. Remote sensing
2. Person recognition
3. One versus all recognition
4. Medical applications.

3.1 Remote Sensing

With the advancement of satellite/sensor generation, the number of remote sensing registers aggregated has doubled in quantity (e.g., terabytes) and elements (e.g., lots of spectral bands). As a result, this field presents surprising and challenging writing algorithm situations. In specifically, classification algorithms should consider:

- Enormous contributions as the examples are gathered over and again for huge spaces [20].
- There are a lot of features because the data is gathered from hundreds of bands.
- A massive number of yields as the classes cover numerous sorts of an area (backwoods, water, farming region) and made articles (streets, houses) [21].
- Missing or defiled information as various groups or satellites might neglect to gather information at specific occasions and ineffectively named (or unlabeled) information as the information should be present handled and doled out on classes.

Hyper-spectral data classification is especially appropriate to classifier ensembles' divide-and-conquer strategy because of these limitations. Before an ensemble puts it all together, multiple classifiers of varying quality can be used to extract the most important information from the raw data. It provides for the sampling of inputs and features to lessen the difficulty each classifier confronts, the estimation of data labels, and the reduction of multi- class problems to numerous two-class problems.

Examples Applications
We will use four different types of ensembles on four different types of remote sensor data to demonstrate the range of applicability of classifier ensembles to remote detecting. The ensembles will range from:

- Random forests and Mountainous Terrain
- Majority voting for agricultural land
- Hierarchical classification of wetlands
- Information fusion for urban areas.

3.2 Person Recognition

Person recognition is the challenge of authenticating a person's identity based on their traits, primarily for security applications. Examples include iris reputation, face notoriety, fingerprint recognition, and behavioral reputation (including speech and

handwriting) to discover a person's characteristics, rather than relying on the unique experience a person might have (including usernames and passwords). To access the account computer, one of the most popular applications for the joint acquisition of knowledge about strategies has been a person's reputation. There are many aspects of individual recognition, some of which make it particularly suitable to be dealt with through shared strategies. A person's reputation can contain two kinds of jobs. In the case of iris, fingerprints, and face popularity, it can frequently use the potentials in laptop vision. Often composed of face popularity, Principal Component Analysis (PCA) perceives several features smaller than the whole picture, despite being more informative. In speech reputation, speech signals' features, including frequencies and amplitudes, must be captured, extracted, or new appropriate capabilities identified. In contrast to more stable features that include fingerprints and faces, speech and other behavioral features can also be represented as a time collection. There are outstanding fault classification fees in personality identification, as in medical and safe anomaly detection programs.

Example Applications

- Unobtrusive Person Identification
- Face Recognition
- Multi-modal Person Recognition
- User-specific Speech Recognition.

3.3 One Versus All Recognition

One versus all recognition incorporates a few distinct issues. One of them is abnormality identification, which is the issue of recognizing surprising examples, i.e., what does not squeeze into the arrangement of distinguished examples. The contrary issue is target acknowledgment discovering what squeezes into a recognized example. Identification of Interruption is an issue that could be addressed the two different ways search for one of a bunch of known kinds of assaults (target acknowledgment) or search for irregularities in the use designs (peculiarity discovery).

When it comes to addressing one versus all recognition issues, ensemble approaches are ideal. In target detection problems, it is easy to imagine having one model for every potential objective and running them all in corresponding to evaluate which model fires or provides a more grounded indicator of acknowledgment than the others. If at least two models fire at almost comparable rates, further processes may be required to distinguish between the various target kinds. An ensemble in anomaly detection can be made up of models that are designed to detect anomalies in a variety of conditions.

As referenced before, the interruption identification issue can be given a role as either an objective location or peculiarity recognition issue. At the point when interruption identification issues are given a role as target location problems, they

are once in a while alluded to as abuse recognition. In these issues, models of realized assaults are concocted, and in the event that current PC framework action coordinates with that depicted by any assault model, it is accepted that the comparing kind of assault is occurring.

Example Applications

Irregularity identification applied to organize interruption recognition, as referenced above, has the issue that any authentic PC framework movement that is not displayed may get unnecessarily hailed an assault. A contributor to the issue is that a considerable lot of the irregularity identification frameworks conceived in the writing are solid models intended to cover all conventions and administrations presented by the PC framework and organization.

- Modular Intrusion Detection
- Hierarchical Intrusion Detection
- Intrusion Detection in Mobile Ad-hoc Networks.

3.4 Medical Applications

There are numerous instances of clinical utilizations of troupe learning. These models incorporate various issues, for example, breaking down x-beam pictures, human genome investigation, and inspecting sets of clinical trial information to search for abnormalities. In any case, the foundation of this load of issues is in evaluating the strength of individuals. This root achieves a few attributes of clinical applications that make them especially troublesome issues. As a rule, such issues have:

Due to the nature of privacy issues, there are few training and testing examples.

- Improper datasets, i.e., a lack of abnormalities or patient instances with a condition.
- An excessive number of ascribes, substantially greater than the preparation and test models, and distinct unclassified costs, with false negatives being far worse than bogus positives.

Because the amount of time spent building and testing models is proportional to the number of patients, such models are typically less expensive than those used in other application regions. The models likewise will in general be moderately imbalanced. Luckily, the quantity of instances of inconsistencies or patients with illnesses is a lot of lower than the quantity of instances of typical patients. Be that as it may, this can present challenges for AI calculations, particularly grouping calculations. Because there are so few positive (i.e., disease-free) instances, classification algorithms are skewed toward anticipating that new cases will be negative. However, mistakenly guessing a negative example (false negative) is almost always worse than predicting a positive example incorrectly. The number of features in human genome analysis and image analysis issues is typically far more than the number of cases. It necessitates the use of feature selection and extraction techniques. The challenge of applying

machine learning to these problems is increased by the need to choose the right method(s).

Example Applications

- Pharmaceutical Molecule Classification
- MRI Classification
- ECG Classification.

4 Conclusion

This research looked at the most common classifier ensemble approaches, along with other statistical combiners such as bagging, stacking, and boosting. Every ensemble technique follows its own set of properties that makes it suitable to specific classifiers and applications. The classifier ensemble community has taken note of four major application fields like medical/biological applications, person recognition, one versus all recognition, and remote sensing. All these applications allow a connected classifier to outperform any single classifier by providing a variety of reasonable-performing but unique base models. This paper assists beginners of ensemble learning technique to understand ensemble learning techniques and their applications in different domains.

References

1. Bishop CM (2006) Pattern recognition and machine learning. Springer, New York
2. Qui J, Wu Q, Ding G, Xu Y, Feng S (2016) A survey of machine learning for big data processing. EURASIP J Adv Sig Process 2016(67):1–16
3. Awad M, Khanna R (2015) Support vector machines for classification, pp 39–66
4. Patel H, Prajapati P (2018) Study and analysis of decision tree based classification algorithms, pp 74–78
5. Pal M (2005) Random forest classifier for remote sensing classification, pp 217–222
6. Rincy TN, Gupta R (2020) Ensemble learning techniques and its efficiency in machine learning: a survey. In: 2nd international conference on data, engineering and applications (IDEA), Bhopal, India, 2020, pp 1–6
7. Utami IT, Sartono B (2014) Comparison of single and ensemble classifiers of support vector machine and classification tree. J Math Sci Appl 2(2):17–20
8. Chen XW, Lin X (2014) Big data deep learning: challenges and perspectives. IEEE Access 2:514–525
9. Freund Y, Schapire RE (1995) A decision-theoretic generalization of on-line learning and an application to boosting. In: European conference on computational learning theory. Springer, Heidelberg, pp 23–37
10. Schapire RE, Singer Y (1998) Improved boosting algorithms using confidence-rated predictions, pp 80–91
11. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29:1189–1232

12. Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining
13. Breiman L (1996) Bagging predictors. Mach learn 24(2):123–140
14. Kuncheva LI (2004) Classifier ensembles for changing environments. In international workshop on multiple classifier systems, pp 1–15. Springer, Berlin, Heidelberg
15. Rodriguez JJ, Kuncheva LI, Alonso CJ (2006) Rotation forest: a new classifier ensemble method. IEEE Trans Pattern Anal Mach Intell 28(10):1619–1630
16. Masoudnia S, Ebrahimpour R (2012) Mixture of experts: a literature survey. Artif Intell Rev 42(2):275–293
17. Jacobs RA, Jordan MI, Nowlan SJ, Hinton GE (1991) Adaptive mixtures of local experts. Neural Comput 3(1):79–87
18. Gutta S, Huang JRJ, Jonathon P, Wechsler H (2000) Mixture of experts for classification of gender, ethnic origin, and pose of human faces. IEEE Trans Neural Netw 11(4):948–960
19. Zhou ZH, Wu J, Tang W (2002) Ensembling neural networks: many could be better than all. Artif Intell 137(1–2):239–263
20. Gislason PA, Sveinsson JR (2006) Decision fusion for the classification of urban remote sensing images, pp 294–300
21. Kumar S, Crawford M (2002) Hierarchical fusion of multiple classifiers for hyperspectral data analysis, pp 210–220

Impact of COVID-19 Pandemic on Indian Stock Market Sectors

M. Saimanasa and Raghunath Reddy

Abstract The start of the COVID-19 pandemic and official lockdown announcements had created uncertainty in global business operations. For the first time, the Indian stock market has significantly impacted. India is one of the most important rising economies in the world and has seen the value of its crucial stock indices plummet by about 40%. There are several studies on the impact of the pandemic on the stock market, but very few studies have focused on a comparative analysis of the first and second COVID-19 pandemic waves. The Fama French model of an event study is used to analyze the response of various sectoral indices during the pandemic. Although all industries were briefly damaged, the financial industry was the hardest hit. Industries such as pharmaceuticals, consumer products, and information technology had favorable or minor effects in both waves. The second wave had an insignificant impact compared to the first one, clearly indicating optimism and normality in the market despite the looming pandemic threat.

Keywords COVID-19 · Nifty sectors · Stock market · Abnormal returns · Event study analysis

1 Introduction

The world is seeing an unrecoverable loss due to the new coronavirus pandemic, which has shaken the global economic and social landscapes. The pandemic's ramifications are devastating, with millions of people sliding into poverty, businesses facing extinction, the workforce enduring job losses, and society undergoing a paradigm shift. The stock market reflects the crisis and its consequences. As a result, the COVID-19 outbreak is crucial for analyzing stock price volatility. COVID-19's effects are not limited to a few fields but are seen in many aspects of society and are

M. Saimanasa (✉) · R. Reddy
CMR Engineering College, Hyderabad, India
e-mail: mannem.manasa254@gmail.com

R. Reddy
e-mail: raghunathreddy@cmrec.ac.in

expected to impact the economy long term. The extent of the influence, however, may differ from industry to sector and country to country.

Due to public worries about diminishing economic activity, less disposable income, and investors' negative attitudes in crisis-like conditions such as pandemics, the stock market's financial performance is predicted to decline. These consequences are reflected in the general market benchmark in the form of diminished liquidity and returns. Sectoral performance, on the other hand, may diverge from the benchmark index due to variances in industry and responsiveness to macroeconomic stimulus. As a result, several industries may exhibit either momentum or contrarian impacts. The study analyzes the premise of falling performance in the stock market in critical sectors in the Indian economy due to the COVID-19 pandemic based on theoretical foundations. Along with the volatility index and the benchmark Nifty index, ten different Nifty sectors were investigated for an abnormal return performance indicator.

The event study methodology is used in investing and accounting to examine stock price volatility and determine whether a single incident might significantly impact the performance of numerous stocks, resulting in unexpected returns. Event analysis is commonly applied to (1) see if markets return responses to economy events and (2) to assess the impact of an activity on a stock owner's wealth, efficiently absorb new understanding. The main difference between event analysis and the autoregressive moving average (ARMA) model is that in time-series data, normality of the data and periodicity are ignored by event analysis. ARMA models, on the other hand, are only relevant to time-series and not to periodic data series. This method also identifies whether the variables are related in a detrimental way.

One of India's leading stock markets, the National Stock Exchange (NSE), is the world's third-biggest stock exchange based on equities trading and boasts the world's largest derivatives market by volume. One of the paper's noteworthy features is its examination of the COVID-19 influence on important Nifty indexes sectors such as financial services, consumer goods, information technology, medicines, and so on (as shown in Fig. 1).

From Fig. 2, it is observed that how Nifty price has affected during COVID-19 pandemic where first and second waves are marked.

2 Literature Review

Several scholars have worked extensively on COVID-19 and its effects on numerous parameters.

The effect of the current pandemic has led to the market crash and also has seen one of the fastest recoveries. The authors in [6] presented that this event provides the unique opportunity to make the right choices and consolidate the market and economy so that the recovery is sustained for the long run rather than short-term market sentiments.

- FINANCIAL SERVICES
- ENERGY
- IT
- CONSUMERGOODS
- AUTOMOBILE
- CONSTRUCTION
- METALS
- PHARMA
- CEMENT & CEMENT PRODUCTS
- TELECOM
- MEDIA & ENTERTAINMENT
- SERVICES0.65%FERTILISERS & PESTICIDES
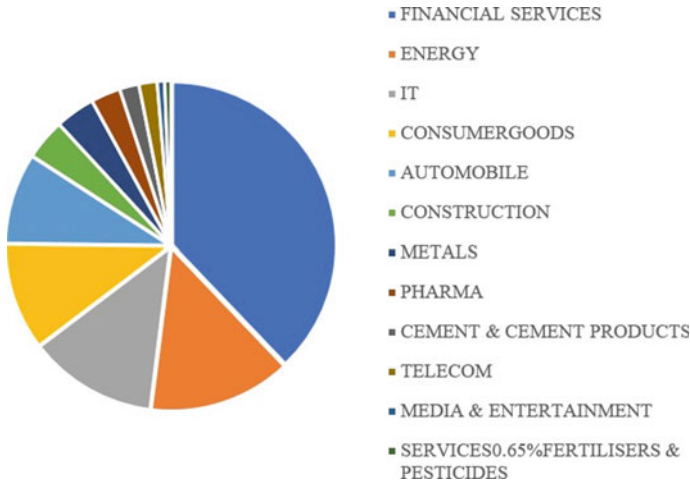
Fig. 1 Weightage of different sectors in Nifty 50 index. *Source* https://dailytrademantra.com/

Fig. 2 Nifty price chart. *Source* Yahoo finance

The article by Barshikar [2] looks at how the pharmaceutical sector adapted itself to meet COVID-19 impact and challenges. They provide new normal impacts of COVID-19.

An event study approach is described for empirically analyzing the market performance and responsiveness of Chinese industries to the COVID-19 pandemic [5]. They found that transportation, mining, electricity and heating, and environment industries were badly affected by the pandemic.

Similar study was carried out in [8] which analyzed how Indian stock markets responded to the pandemic. The authors examined the short-term impact of the pandemic on the Nifty index and its constituent sectors. They employed three models: the constant return model, the market model, and the market-adjusted model. They

stated that all sectors were briefly disrupted. The medical, consumer items, and IT sectors all had favorable or minor effects.

In paper [9], the authors observe the outbreaks and conclude that markets are adversely affected in the short-run but eventually in the long run, and they improve their performance. They focus on the travel industry, technology, entertainment, and gold as the potential areas for the study.

COVID-19 is the first global environmental catastrophe of the twenty-first century. The impact of COVID-19 on the Indian stock market during various lockdown periods is examined [1]. A comparison of outperformance and sustainable development in a number of South Asian countries is also included in the research.

The authors of this research [3] examine the potential impacts of the coronavirus "COVID-19" on the stock market and then offer strategies for individuals to benefit from a market disrupted by a worldwide viral outbreak. The authors recommend shorting industries that will be hit hard by the virus in the short term, then buying back in after the price has dropped dramatically.

Article [7] explained how financial markets are very volatile and continue to fluctuate based on the company's performance, historical data, market value, and news and timeliness.

The paper by Fama and French [4] explains the common risk factors that affect returns on stocks and bonds. Fama and French study explain market, size, leverage, and book-to-market equity in stock returns.

3 Methodology

3.1 Data

The research investigates how the COVID-19 epidemic affected the Indian stock exchange in both the first and second waves of the virus. For this analysis, 10 days pre and 20 days' post-event data is used. The influence of COVID-19 on Indian stock market returns is investigated using an event study technique and the Fama and French three-factor model. Stock prices were sourced from investing.com, which were adjusted for stock splits and dividends. The research took into account daily closing prices. To analyze, the was data collected from November 01, 2019 to February 24, 2022. The closing prices of 10 sectors (automobile, bank, Indian consumption, finance, metal, media, IT, energy, reality, and pharma) are collected from the investing.com database.

3.2 Proposed Methodology

The impact of an event is determined by calculating abnormal returns. The event research approach aids in determining if unforeseen occurrences resulted in abnormal returns. The related abnormal returns are calculated using a Fama French model. The Fama French model outperforms the CAPM model in predicting variance in excess return over Rf.

The Fama French three-factor model is a variant of the CAPM model. Three components are used in the Fama French three models are market risk, the outperformance of small-cap companies relative to large-cap companies, and the outperformance of high book-to-market value companies versus low book-to-market value companies. Eugene Fama and Kenneth French of the University of Chicago devised the Fama French three-factor model [3].

Formula for Fama French three-factor model

$$r = \text{rf} + \beta 1(\text{rm} - \text{rf}) + \beta 2(\text{SMB}) + \beta 3(\text{HML}) + \epsilon \tag{1}$$

r	Expected rate of return
rf	Risk-free rate
β	Factor's coefficient (sensitivity)
$(\text{rm} - \text{rf})$	The market risk premium is defined as the difference between the market's predicted return and the risk-free rate. It compensates an investor for the greater volatility of returns over and beyond the risk-free rate by providing an excess return.
Small Minus Big (SMB)	Small-cap firms have historically outperformed large-cap corporations in terms of returns. SMB is a size impact depending on a company's market capitalization. SMB is a statistic that measures small-cap companies' historical advantage over large-cap firms. After identifying SMB, the beta coefficient may be calculated using linear regression.
High Minus Low (HML)	Value stocks' historical excess returns (high book- to-price ratio) stocks with a lot of growth (low book-to-price ratio). The HML premium is a value premium. The HML factor's beta coefficient may be obtained using linear regression, just like the SMB factor.
ϵ	Risk.

Step1: Defining the event window

It is the first stage in event analysis. On the evening of March 24, 2020, the Indian government ordered a statewide lockdown. 2020-03-24 as the first event and 2021-03-23 as the second event. The anticipation window and the adjustment window are two parts of the event window. The anticipation window is 10 days previous to the event and 20 days to post-event following the event. The virus continues to startle
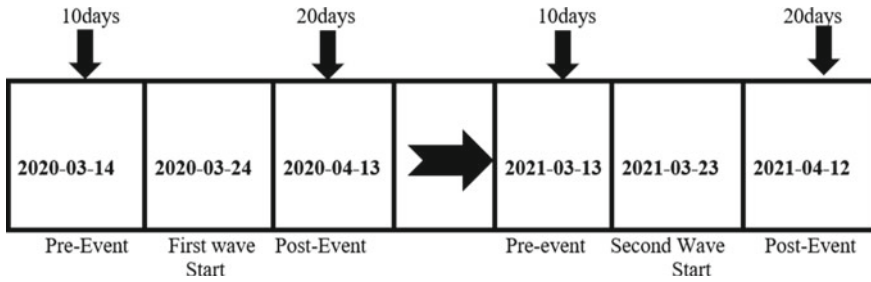
Fig. 3 Event window for two pandemic waves

people on a daily basis and is continuously being detected throughout the world. Figure 3 shows the event window for two pandemic Waves.

Step 2: Calculating abnormal returns

Over a particular time period, an abnormal return reflects the exceptionally significant profits or losses achieved by a certain investment or portfolio. Fama French model is used for calculating abnormal returns.

4 Empirical Results

In this study, abnormal returns were obtained from 10 different sectors and plotted graphs for each sector for two events, and the axes taken are window period and abnormal returns. From the graphs plotted, it is observed that the results are mixed with positive and negative returns in both the waves. Only a few sectors have stable results.

During the COVID-19 epidemic, the Nifty pharma index saw a lot of volatility. During first wave (Fig. 4), the Nifty pharma index increased almost 20%, compared to a 3.5% drop in the Nifty. The Nifty pharma index increased 61% in 2020, surpassing the Nifty's 15% increase. However, it declined 6% and 2% in January and February. The index has risen over 9% in the second wave (Fig. 5), while the Nifty has lost roughly 3.5%. The pharmaceutical business is committed to safeguarding human safety. India is the world's largest supplier of generic medications, accounting for 20% of global exports in this category.

As several indices nearly doubled during this time, some stocks turned out to be multi-baggers. The Nifty metal index is intended to reflect the performance and behavior of India's metals industry, which includes mining. Metal prices have fallen following a record-breaking rise, causing the Nifty metal Index (NIFTYMET) to fall 3.8% in the first wave (Fig. 6). The stock market has been dominated by the COVID-19 pandemic for the past year, resulting in massive returns. Nifty metal outperformed the other indexes in the second wave (Fig. 7).

Fig. 4 Pharma: COVID-19 first wave
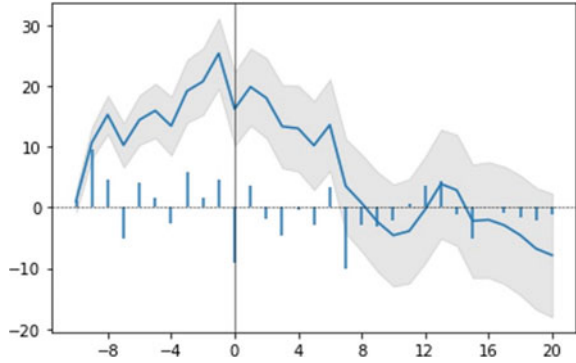
Fig. 5 Pharma: COVID-19 second wave
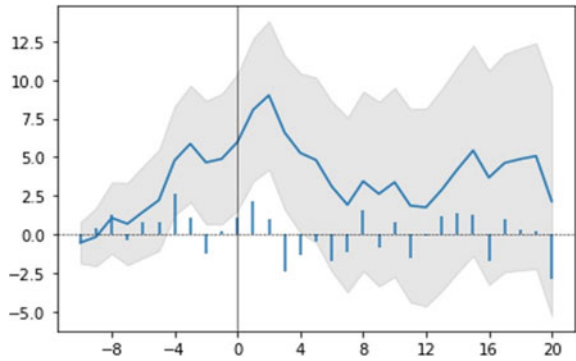
Fig. 6 Metal: COVID-19 first wave
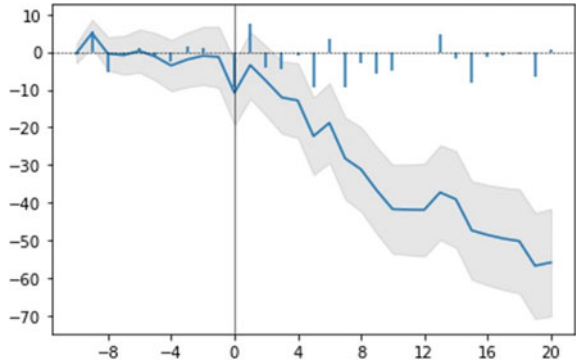
Auto stocks were under pressure as investors hedged their bets as the second wave of the coronavirus pandemic reached India, resurrecting the threat of a shutdown. Compared to the first wave (Fig. 8), the second wave has positive returns (Fig. 9). As sales improve, automakers are quickly recovering from the impact of a severe shutdown imposed in 2020. The index has touched high and low returns.

Fig. 7 Metal: COVID-19 second wave
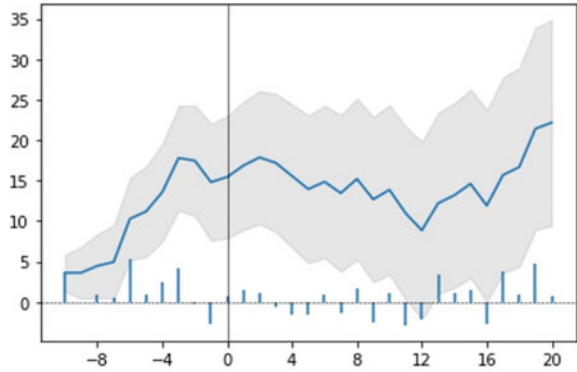
Fig. 8 Auto: COVID-19 first wave
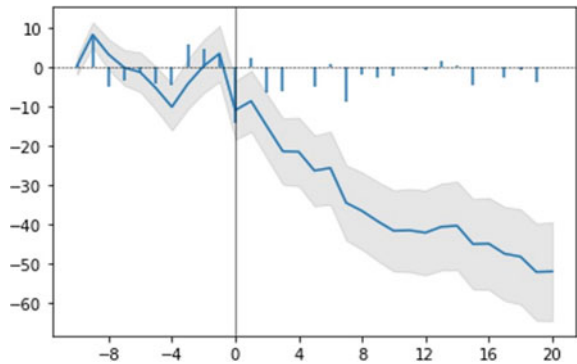
Fig. 9 Auto: COVID-19 second wave
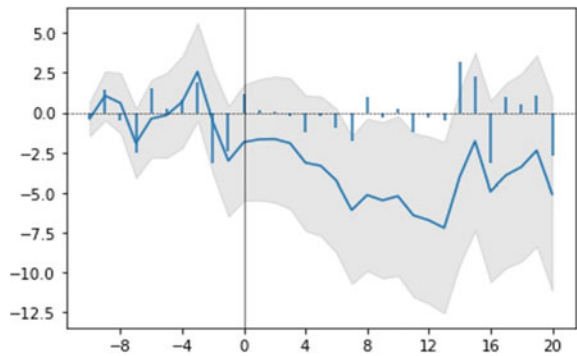
IT sector fell from 0.8 to 1.2% in the first wave (Fig. 10); during phase 1, it was in negative correlation. But it raised quickly in the second wave (Fig. 11). Nifty IT was in positive correlation. This study also analyzed the remaining sector, which found mixed results in both the waves but found only pharma, metals, and IT sector had positive, and finance was hit big loss during both the waves.
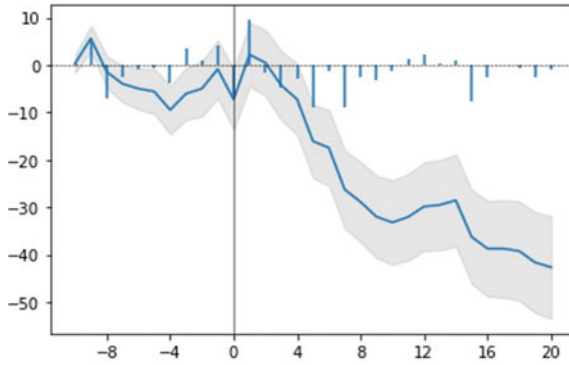
Fig. 10 IT: COVID-19 first wave
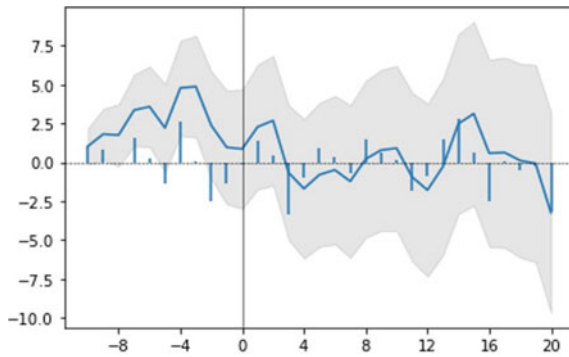
Fig. 11 IT: COVID-19 second wave

Figure 12 observed that only the Nifty pharma sector has positive results compared to the remaining 9 sectors in the COVID-19 first wave.

Figure 13 observed that in the second wave, along with the Nifty pharma sector Nifty metal sector also had positive results that were higher than the pharma sector.

5 Conclusion

The influence of COVID-19 on the Indian stock market was investigated in this study, which looked for abnormal returns at the pandemic's beginning. We use event study methodology, including the Fama French three-factor model. Abnormal returns were seen for many days before and after the event. On most days following the declaration of total lockdown, all models exhibited consistently favorable AARs. Pandemic was shown to have increased stock market risk in general. However, the findings are diverse and substantially impacted by the industries. All industries were briefly
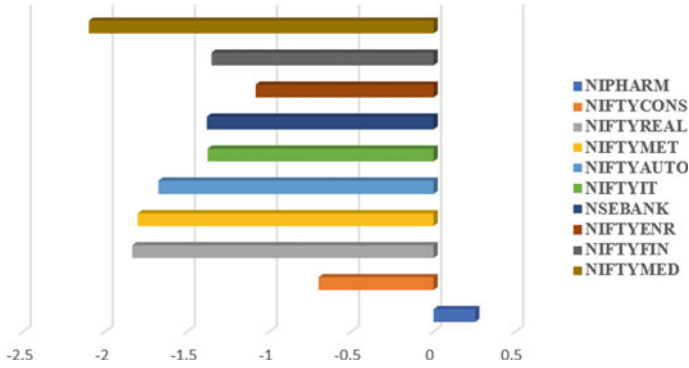
Fig. 12 Abnormal mean returns for the 10 sectors in COVID-19 first wave
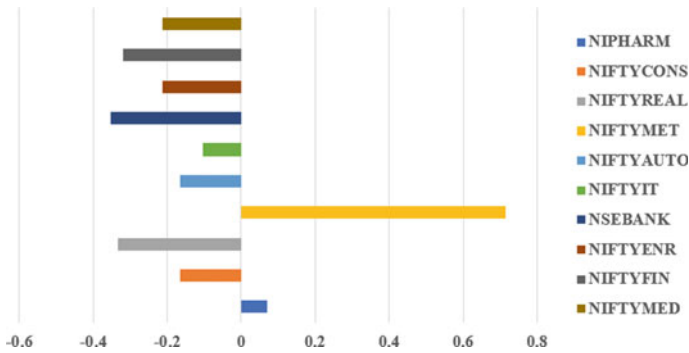
Fig. 13 Abnormal mean returns for the 10 sectors in COVID-19 second wave

damaged, but the financial industry was the worst hit. Pharma, metals, and information technology all had favorable or minor effects. The analysis found that NSE-listed companies negatively reacted to the COVID-19 outbreak, with over 6% of negative AR event windows. The price response of several sectors to the pandemic outbreak is then examined. In event windows, the severely negatively impacted sectors have shown a negative abnormal return of more than 10%, including financial sectors, metal, automobiles, etc. This research also contributes to a better understanding of how different industries respond to the pandemic. Furthermore, because the size of a firm is a significant driver in absorbing the impact of extreme occurrences, this research can help us understand how the pandemic would affect businesses of various sizes.

References

1. Ahmed F, Syed AA, Kamal MA, de las Nieves López-García M, Ramos-Requena JP, Gupta S (2021) Assessing the impact of covid-19 pandemic on the stock and commodity markets performance and sustainability: a comparative analysis of South Asian countries. Sustainability 13(10):5669
2. Barshikar R (2020) Covid 19—impact and new normal for pharmaceutical industry (Part–I). J Generic Med 16(3):112–119
3. Fama EF, Fisher L, Jensen M, Roll R (1969) The adjustment of stock prices to new information. Int Econ Rev 10(1)
4. Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. J Financ Econ 33(1):3–56
5. He P, Sun Y, Zhang Y, Li T (2020) Covid-19's impact on stock prices across different sectors—an event study based on the Chinese stock market. Emerg Mark Financ Trade 56(10):2198–2212
6. Sarkar S (2020) Covid-19 impact on stock market and economy of India. Int J Res Eng Sci Manag 3(11):55–64
7. Singh A, Gupta P, Thakur N (2021) An empirical research and comprehensive analysis of stock market prediction using machine learning and deep learning techniques. In: IOP conference series: materials science and engineering, vol 1022. IOP Publishing, p 012098
8. Varma Y, Venkataramani R, Kayal P, Maiti M (2021) Short-term impact of covid-19 on Indian stock market. J Risk Financ Manag 14(11):558
9. Yan B, Stuart L, Tu A, Zhang T (2020) Analysis of the effect of covid-19 on the stock market and investing strategies. Available at SSRN 3563380

Advance Warning and Alert System for Detecting Lightning Risk to Reduce Human Disaster Using AIoT Platform—A Proposed Model to Support Rural India

Ome Nerella and Syed Musthak Ahmed

Abstract Lightning is one of the serious natural calamities, which has led to several deaths in our country. Around 2000 people lost their lives every year due to the lightning risk in India. According to Annual Lightning Report 2020–2021, the lightning deaths recorded 96% in rural areas and 4% in urban areas. The primary causes of lightning deaths are mainly due to standing under tree, indirect hit, and direct hit. Present lightning alert systems are based on satellite and radar systems, which are expensive due to the processing of the lightning data. Existing lightning alert information is not reaching to the rural areas due to low literacy, less usage of smart phones and also due to lack of awareness about lightning effects and causes and also lack of knowledge on alert apps. Therefore, there is a need to bring awareness among rural population about the effects of lightning disasters and/or alerting them of the consequences. The proposed paper presents a method to predict the lightning and alert them to overcome disaster. Here, artificial intelligence (AI) along with Internet of Things (IoT) is incorporated to alert the rural people in advance about lightning by various means such as announcement system, voice calls, alerting by sounding horn, WhatsApp messages, and mobile apps. The proposed system takes real-time weather and lightning data from many weather stations into account, and it has been trained to predict lightning situations using machine learning algorithm and making alert to the rural areas.

Keywords AI (Artificial intelligence) · AIoT (Artificial intelligence of things) · IoT (Internet of things) · Lightning detection · LoRa (Long range) · Weather stations

O. Nerella (✉) · S. M. Ahmed
School of Engineering, SR University, Warangal (Urban), Telangana, India
e-mail: omenerella@gmail.com

S. M. Ahmed
e-mail: musthak.ahmed@sru.edu.in

1 Introduction

Lightning is a natural disaster, which cannot be eliminated but can be predicted and prevented further from deaths due to lightning. The intensity of lightning varies from location to location with coastal areas recording more lightning strikes as compared to other zones. According to the Annual Lightning Report 2020–2021, Bihar has the largest number of lightning deaths, while Tamil Nadu has the lowest number of lightning deaths. Odisha has the highest number of lightning strikes, while Naga-land has the lowest number of lightning strikes. By implementing early lightning warning system, 70% of deaths in Odisha and Andhra Pradesh have got reduced. But, populations in tribal areas are more vulnerable to lightning effects due to their tinned shelters.

Figure 1 shows the statistical information of lightning effects giving percentage of data at various zones, areas, places, locations, etc. Figure 2 gives the information on primary causes of deaths due to lightning.

Depending on the formation of clouds, there are three forms of lightning: in-cloud, cloud-cloud, and cloud-ground. Lightning can be detected by three fundamental methods: acoustic, optical, and electric field. Acoustic lightning detection refers to listening thunder to identify lightning. Optical lightning detection refers to sensing
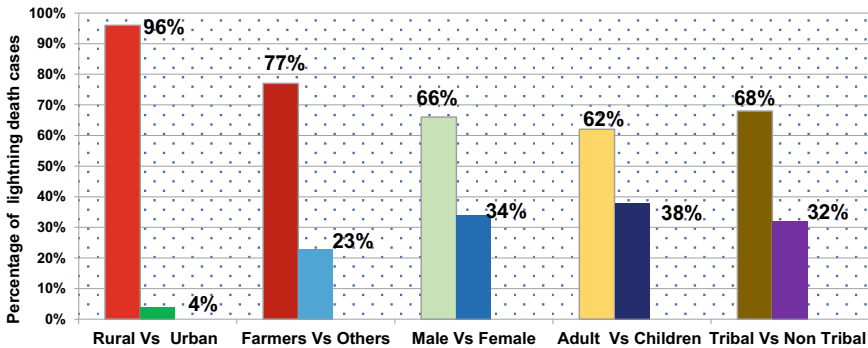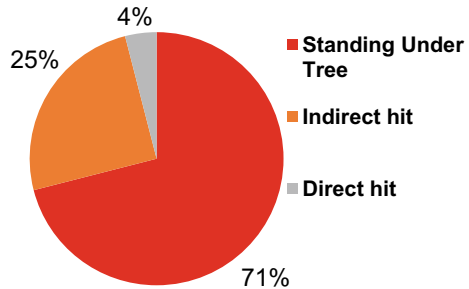

Fig. 1 Lightning death cases in various sectors

Fig. 2 Primary causes of lightning deaths

lightning by space-based global navigation satellite technique. Electric fields method of detection refers to the way of detecting electric field or radio wave using ground-based system. However, ground-based systems are more preferred over space-based systems.

2 Literature Review

Real-time weather and lightning forecasting are critical for warning people ahead of time about natural disasters and reducing the loss of life and property damage. Kanchi et al. [1] created a lightning detection and alert system that uses cloud computing technology to alert people about lightning through mobile apps. This alert system was created with the use of a lightning detector, a microcontroller, and an open-source cloud platform. This system developed by them can collect storm and lightning strike data over a distance of up to 40 km from the deployed location.

Bobbadi et al. [2] in on smart phones and warn users of lightning strikes in real time based on location as well as the location of the hit. The main goal of this app is to reduce deaths caused due to thunderstorm by sending early warnings to users. This app provides the location of the storm even in the absence of Internet.

Kotal et al. [3] in Jharkhand created "A location-specific newscast and Short Message Service dispersion system for rainfall and lightning alert" in Jharkhand, India. They presented a new type of spot newscast and message service dispersion system that warns around 3 h ahead of about rainfall and lightning warning across Jharkhand. They incorporated mosaic composite reflectivity data from a variety of radars to pinpoint the emergence and tracking of thunder cells. Thunder cells' future location, as well as their expansion or dissipation as a function of radius, are predicted using extrapolation based on past trends. Second, the Jharkhand Space Applications Center's Web-based GIS site generates a text message for the predicted area, which is to provide the information to the proper governmental officials and mobile operators in the region. Ultimately, all active mobile users in that region are issued a warning SMS at that moment.

The use of an atmospheric electric field device for lightning detection is common. The lightning alarm is triggered by the electric field threshold. The best threshold may be determined by examining the efficiency of lightning early warning because the threshold value has a direct impact on the warning's accuracy. Lightning surveillance is required to minimize the economic damage incurred by lightning incidents. Many lightning monitoring systems exist today, including lightning locating radar, fore-casting satellites, and environmental electric field instruments [4]. By observing the associated lightning signal, these instruments can identify or anticipate the appear-ance of lightning. These pieces of equipment have varied costs and effects. Radar data and forecast satellite have higher costs. The lightning localization system can only record the presence of lightning, not anticipate it. Lightning occurs when the environ-mental potential gradient approaches the climate's breakdown electrical potential.

As a result, the variation in environmental electromagnetic field can be used to monitoring lightning whenever lightning strikes, the amplitude of the electric field in the environment will show as a random vibration.

A system was created by Mostajabi [5]; it recognizes lightning conditions using a machine learning algorithm and informs individuals 30 min advance within a 30 km area of the deployed place. It can even cover isolated areas that may be out of radar and satellite coverage and have no network communication. Air pressure, ambient temperature, humidity, and wind velocity are the four variables that were considered. The metrics were compared to data from lightning detection and positioning systems and algorithm learned the conditions under which lightning occurs using this way. According to the researchers, after training, the system provided predictions that were over 80% correct.

Thunderstorm prediction is said to be the most difficult subject in weather forecast owing to the restricted temporal and geographical extension, either dynamically or physically, that lightning causes more deaths each year than tornadoes. Kalbende and Shelke [6] suggested a method that employs a hybrid model to get the best outcome for detecting not whether lightning the earth. In the identification of thunderstorms, the resulting mechanism surpasses previous strategies such as MOM, LM, QKP, STP, CG, and DBD models. Accuracy is determined by computing the four essential performance metrics of specificity, accuracy, precision, and sensitivity. The proposed system is used by the meteorology department to identify thunderstorm and decide whether or not they are life threatening.

Each year, lightning kills at least 25,000 people around the world. Hwang et al. [7] proposed prototyping an RF signal-based lightning alarm system employing Internet of Things (IoT) connectivity. Using a lightning detector, the researchers hope to explore the induced voltage caused by lightning strikes.

This device was initially developed to detect lightning using a radio frequency receiver with a frequency range of 300–500 kHz.

If thunder creates a sufficient induced voltage within 10 km of the detector, the lightning detector may be triggered. This detector saves the produced EMF and time lag observed lightning discharge, as well as the noises, for further analysis. The accuracy of the lightning detector was judged to be 88.33%. A cloud system supplied the analyzed data to the smartphone application. The induced voltage is inversely proportional to the lightning striking distance.

$$D = 0.8557V^{-1.396} \tag{1}$$

In Eq. 1, D is distance of lightning strike, and V is the induced voltage. Thus, the closer a lightning strike is detected, the higher the induced voltage. Ground-based and space-based approaches are two major technologies for detecting lightning in real time [8]. Ground-based lightning location systems employ time of delivery and route knowing approaches, whereas satellite-based lightning location systems use optical imaging and direction finding in VHF range. The TOA approach simultaneously analyzes the start time of a specific characteristic of a lightning event's electromagnetic waveform at many sensors. If a lightning strike happens, the time,

latitude, and longitude are calculated. In the VLF-MF frequencies, magnetic direction finding is often used, whereas interferometer is commonly used in the VHF frequency. Satellite-based LLS operate in the IR/optical frequency range (30–300 THz), whereas ground-based LLS operate in the VLF-VHF frequency range (3 Hz–300 MHz).

An integrated lightning strike protective (LSP) layer made of 3D printing technology nickel-plated carbon fiber mesh is proposed by Ming et al. [9]. The electrical and thermal conductivity of the fibers is ensured by the conductive nickel coating. Low population density, max strength, and stiffness are all appealing characteristics of continuous carbon fibers. Because of 3D printing technology, it is now possible to produce LSP layers at a low cost for complicated and superstructures. The nickel-plated carbon fiber mesh successfully dispersed arcs and suppressed stresses. As a result, the damaging area and depth inside the surface layer are minimized. These results suggest that the suggested approach might be useful in future aerospace applications.

Thunderstorms over India are known for their catastrophic lightning activity. As a result, it is crucial to comprehend the long-term variations in lightning incidence and intensity, as well as how they interact with diverse causative elements. The most abundant lightning incidences are found along the Foothills of the Himalayas, Indo-Gangetic plains, and coastal areas, with the intensity of these lightning strikes being strongest along the coastal areas and Bay of Bengal, according to long-term TRMM satellite-based lightning observations (1998–2014).

The primary determinants for lightning events are convective available potential energy, aerosol optical depth, and total column water vapor [10]. Lightning radiances are brightest along the coast and around the oceans due to the influence of hydrometer strength on the thunder electrostatic potential formula caused by increased moisture level in these areas. The Himalayan hills, followed by coastal areas and the Indo-Gangetic flats, have the greatest global weather average of lightning strikes that can be linked to atmospheric circulation convection, wetness infiltration, and aerosol thunderstorm interaction.

Between 1998 and 2014, the frequency of lightning increased by 1–2% per year across all Indian areas, with significant regional variation.

Machine learning at the edge for AI-enabled IoT devices was reviewed by Merenda et al. [11]. Machine learning on IoT devices minimizes congestion issues by permitting operations to be done near to sources of data, maintaining data privacy while uploading and lowering battery usage for enabling continuous radio links to cloud servers or gateways. The purpose of this study was to give a high-level overview of the essential strategies for assuring the application of machine learning techniques in the IoT framework on constrained hardware, paving the way for the network of aware objects.

The fuzzy logic technique was used by Nyap et al. [12] to construct a lightning prediction algorithm. When temperature, humidity, and dew point temperature all reach a certain level, the system may anticipate lightning. Temperature, dew point temperature, and humidity are used to predict the lightning. The goal of this study is to build and construct a lightning prediction system that uses fuzzy logic and alerts people when lightning is approaching. After being evaluated with genuine data from

the meteorological department, it was possible to establish that this system's accuracy in lightning prediction is greater than 95%.

Verma et al. [13] used IoT with a linear regression machine learning model to develop a weather prediction system. It makes use of the NodeMCU, a digital humidity and temperature sensor, a light intensity sensor, the ThingSpeak cloud, and a Jupyter notebook. The data from the sensors are analyzed by the MCU, which then sends it to an open-source cloud and a Google spreadsheet. The data from the deployed location are fetched by a machine learning algorithm that has been taught to predict weather conditions in a Jupyter notebook environment. This system has a greater prediction accuracy (84%) when compared to artificial neural network machine learning models and decision tree algorithm.

Sharma and Prakash [14] developed a real-time climate tracking system based on IoT. It is made up of different weather sensors and a NodeMCU that is stationed in Gorakhpur. The NodeMCU was turned into an embedded Web server that served Webpages depending on user requests using embedded C, HTML, and HTTP. Everyone may check the weather's state from anywhere using a Web server, without relying on any application or Web site. The data are open to the public, and this technique is used to monitor the weather in Gorakhpur.

Raval et al. [15] created a system that uses IoT to monitor numerous character-istics of weather and uses an auto-regression integrated moving average machine learning model to predict future values. It uses DHT11 and BMP 280 sensors to detect temperature and humidity, as well as pressure and altitude. With the help of the Internet, NodeMCU processes the sensor data and sends the information to the ThingSpeak open-source cloud. The data are collected from the cloud storage and fed into a time series analysis method called auto-regression integrated moving average, which predicts and visualizes patterns and trends in the data.

3 Existing Systems

There are 3 basic methods available to predict lightning and alert people: Geospatial-based system, SMS-based system, and IoT-based systems. The block diagram representation of these systems shown in Figs. 3, 4, and 5.

3.1 Geospatial-Based Lightning Alert System

The block diagram of such system is depicted in Fig. 3. This type of system communicates with the Earth network to collect lightning data. The Vajrapat mobile app is developed for this purpose to help the people and make them alert while they are approaching a lightning zone. This type of lightning alert system is being installed in the state of Andhra Pradesh.
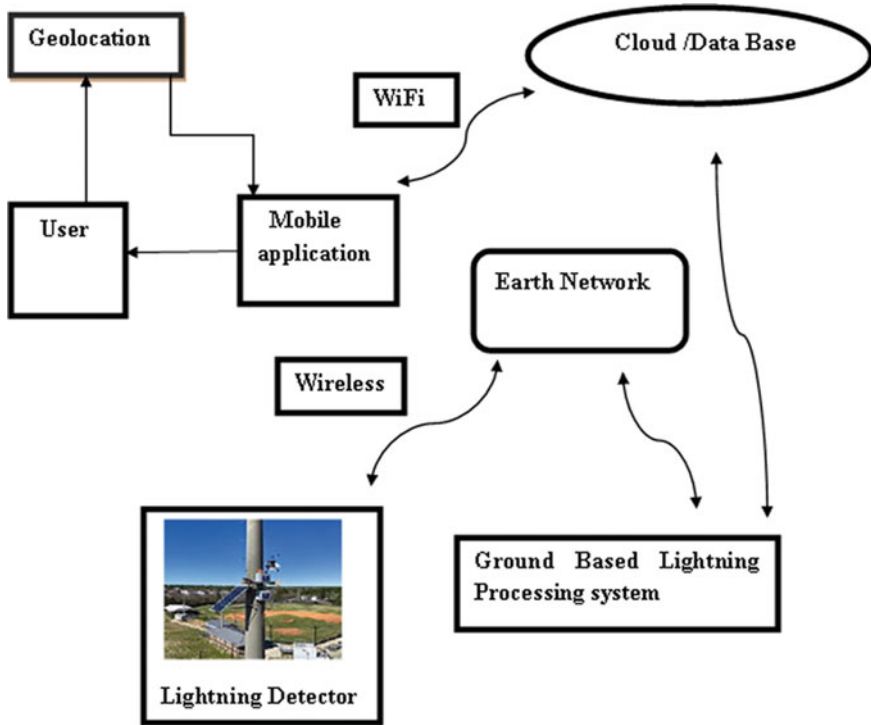
Fig. 3 Geospatial-based lightning alert system
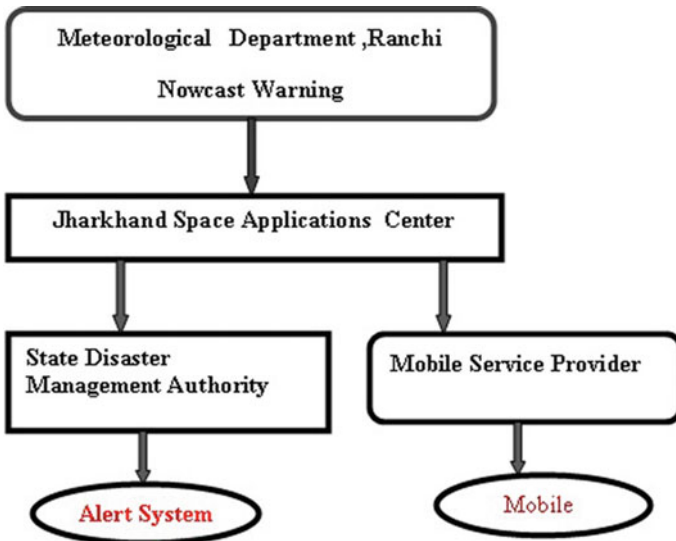
Fig. 4 SMS-based lightning alert system
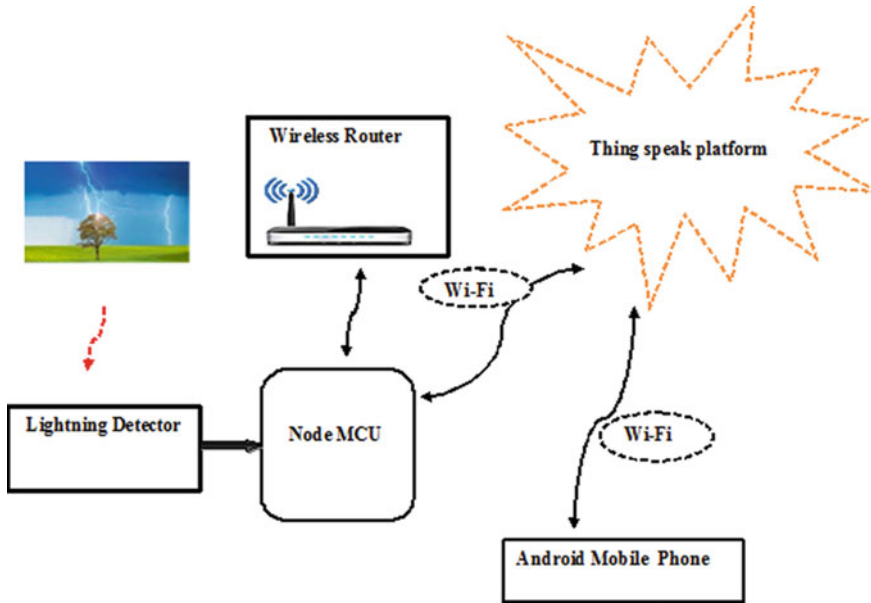
Fig. 5 IoT-based lightning alert system

3.2 SMS-Based Lightning Alert System

Figure 4 shows the SMS-based lightning alert system. This system collects the data about weather from satellite/radar. The Meteorological Department analyzes the climate changes and sends the lightning alert information to the State Disaster Management Authority (SDMA) and mobile service providers in advance to notify the people about lightning storms. This type of lightning alert system is being installed in the state of Jharkhand.

3.3 IoT-Based Lightning Alert System

Figure 5 depicts the IoT-based lightning alert system. This system is developed using NodeMCU, AS3935 lightning detector, and ThingSpeak platform. This system tracks meteorological data in a region and makes the data available for the user having Internet facility.

4 Proposed System

The proposed system incorporates intelligence into the existing IoT system without any human intervention. This has led to development of new system, i.e., Artificial Intelligence of Things (AIoT). This model has the benefits of enhanced statistics, information management, IoT functions, and machine to human communication.

The proposed system is shown in Fig. 6. This system incorporates weather-based stations, a gateway (AIoT edge computing device), a database, a cloud platform, and a client system. This system utilizes LoRa along with AIoT technology. The data are processed and analyzed in real time near the sensor node using edge computing. The proposed system communicates with the available weather stations to get the real-time data on weather and lightning. The technology in it uses machine learning algorithm to predict the lightning and sends this information to a cloud platform and alert system. Later, various other methods are included to warn people in rural areas about lightning by the way of voice calls, announcements, and sound horn besides WhatsApp messages and mobile apps.

The long-range extension is achieved by LoRa technology as shown in Fig. 7. Here, a microcontroller unit, lightning detector, various weather sensors, GPS module, actuators, LoRa module, Wi-Fi module, and solar panel are utilized. The MCU analyzes the sensor data and delivers weather and lightning data to the gateway (AIoT Edge Computing Device) through LoRa module, and the process is continued as explain by Fig. 6. The suggested system does not require Internet access at the deployment location since it uses LoRa and AIoT technologies, which mitigates the disadvantages of existing lightning alert systems.
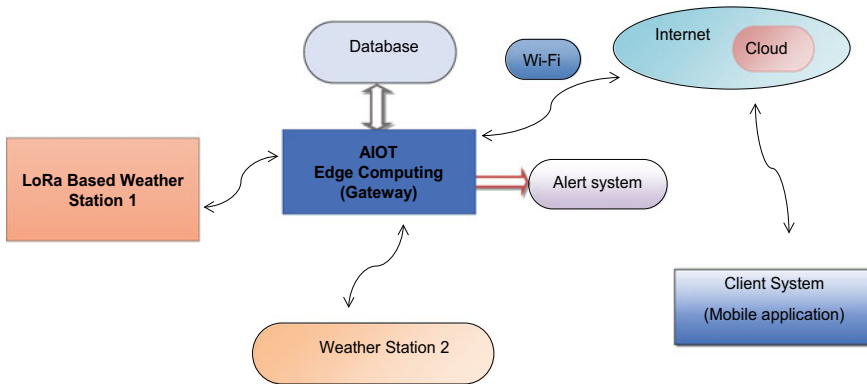
Fig. 6 Proposed AIoT-based lightning prediction and alert system for rural India
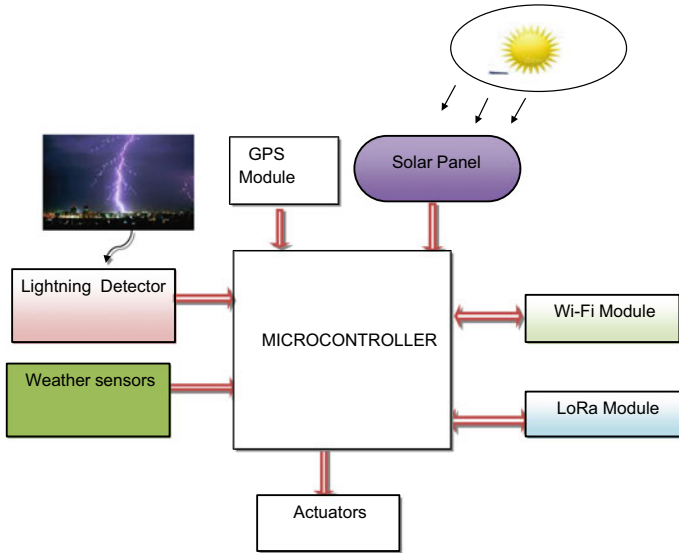
Fig. 7 LoRa-based weather station

5 Conclusion

In the present work, ground and satellite-based methods of detecting the lightning and alerting the people prior to the disaster are bought out. The setbacks of these systems are also enlightened. Contrast to the existing method, the present method as added advantages with respect to cost and processing of lightning data. The various meteorological parameters such as atmospheric temperature, pressure, wind speed, humidity, and atmospheric electric field are made use of the predictors of lightning. This system can be deployed in remote and rural areas where Internet is unavailable as much of Indian rural people do not have the necessary education of disaster due to lightning, use of mobile apps and smart phones. Machine learning algorithm is developed, and the hardware is tested with various sensors to predict the weather parameters and lightning condition accurately. The developed system will be added advantages to the rural people and a new low-cost lightning alert system for rural India to minimize the catastrophe.

References

1. Kanchi RR, Palle D, Sreeramula VP (2018) Design and development of IoT-cloud-based lightning/storm detection system with an SMS alert on android mobile. In: International conference on recent innovations in electrical, electronics & communication engineering, pp 2301–2305. http://doi.org/10.1109/ICRIEECE44171.2018.9008617

2. Bobbadi C, Anudeep DSVNSS, Vasavi S, Harikiran C (2020) Geospatial based thunder storm alert system. Int J Grid Distrib Comput 13:2764–2771

3. Kotal SD, Roy SS, Sharma RS, Mohapatra M (2021) A location-specific nowcast and SMS-based dissemination system for thunderstorm and lightning warning over Jharkhand, India. J Curr Sci 120:1194–1201. http://doi.org/10.18520/cs/v120/i7/1194-1201

4. Chen Y, Tao H, Gu SQ (2020) Optimization of lightning warning effectiveness assessment method for atmospheric electric field. In: 4th IEEE conference on energy internet and energy system integration, pp 1149–1153. http://doi.org/10.1109/EI250167.2020.9346570

5. Mostajabi A (2019) Artificial intelligence can now predict when lightning will strike. J Clim Atmos Sci 2:1–15

6. Kalbende R, Shelke N (2017) Optimized thunderstorm and lightning detection system. Int J Sci Res (IJSR) 6:840–844

7. Hwang YJ, Wooi CL, Rohani MNK, Mehranzamir K, Arshad SNM, Ahmad NAY (2020) Prototyping a RF signal-based lightning warning device using with Internet of Things (IoT) integration. J Phys 1432:1–7. http://doi.org/10.1088/1742-6596/1432/1/012078

8. Nag A, Murphy MJ, Schulz W, Cummins KL (2015) Lightning locating systems: insights on characteristics and validation techniques. J Earth Space Sci 2:65–93. http://doi.org/10.1002/2014EA000051

9. Ming Y, Xin Z, Zhu Y, Zhang C, Yao X, Sun J, Wanga B, Duan Y (2021) 3D printed nickel-plated carbon fiber mesh for lightning strike protection. J Mater Lett 294:1–4

10. Chakraborty R, Chakraborty A, Basha G, Ratnam MV (2021) Lightning occurrences and intensity over the Indian region: long-term trends and future projections. J Atmos Chem Phys 21:11161–11177. http://doi.org/10.5194/acp-21-11161-2021

11. Merenda M, Porcaro C, Iero D (2020) Edge machine learning for AI-enabled IoT devices: a review. J Sens 20:1–34

12. Nyap LCC, Waheeb W, Luokse J (2020) Lightning prediction using fuzzy logic technique. J Appl Technol Innov 4:1–9

13. Verma G, Mittal P, Farheen S (2020) Real time weather prediction system using IoT and machine learning. In: 6th international conference on signal processing and communication (ICSC-2020), pp 322–324. http://doi.org/10.1109/ICSC48311.2020.9182766

14. Sharma P, Prakash S (2021) Real time weather monitoring system using IoT. In: ITM (Information technology, computer science and mathematics) web of conferences, vol 40, no 01006, pp 1–10. http://doi.org/10.1051/itmconf/20214001006

15. Raval MP, Bharmal SR, Hitawla FAA, Gupta P (2020) Machine learning for weather prediction and forecasting for local weather station using IoT. Int Res J Eng Technol (IRJET) 07:419–423

Fetal Health Prediction Using Machine Learning Approach

C. Chandana, P. N. Neha, S. M. Nisarga, P. Thanvi, and C. Balarengadurai

Abstract Fetal health is extremely important for the good health of the fetus, so it is very necessary to notify the doctor about the uterus of pregnant women. Prenatal diagnosis including screening and diagnostic tests can provide valuable information about the health of fetus and to understand the risk. The objective of this study is to provide the convenient services to pregnant women and clinicians by collecting the secondary data from various sources. We use SVM algorithm which is associated with ML-based conventional algorithms to predict the fetal health, through which we can know whether the fetal status is normal or abnormal by considering the health parameters of fetus. By using this approach, we can promote healthy administration and development of the fetus. As result shows, this method can be more beneficent for early detection of well-being of fetus as well as pregnant women.

Keywords AI · Fetal · ML · Prediction · SVM

1 Introduction

Health care is extremely important affecting lives of people worldwide. Nowadays, various information technologies to access healthcare services remotely and manage the health have become the crucial part of healthcare system. In pregnancy, a health-care provider will want to check the health of the unborn baby (fetus). Discovering reliable information of pregnant women is significant to decrease high rate of deaths or any other abnormalities in the fetus. Complexities during the period of pregnancy lead fetus to very serious problems which limits the proper growth of fetus or demise. Hypertension is the most repeated medical concern at the time of gestation. Decrease in the fetal mortality is correlated with a strong depletion in stillbirths. So that the early examination of fetal health can be determined based on so many parameters such as fetal growth, fetal weight, baseline values, acceleration, fetal moment, uterine

C. Chandana · P. N. Neha · S. M. Nisarga · P. Thanvi · C. Balarengadurai (✉)
Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysore, India
e-mail: balarengadurai@vvce.ac.in

contractions and so on. Ultrasound scanning has become a highly suggested examination in prenatal diagnosis in many countries. Diagnostic methods are based on parent characteristics and transvaginal ultrasound. Fetal movement counting can also provide valuable information. Fetal ECG can be recorded by the detection of amplitude and shape changes which occurs in the abdominal region. Variability of the heart rate is acknowledged to approach the autonomic nervous system. It can be quantified through signal processing algorithms for heart rate frequency. Fetal state monitoring is being examined by UCI CTG dataset which is used to speculate the classification purpose. For selecting classifier that constructs an accurate learning model for forecasting the risk based on CTG dataset, data mining and machine learning algorithm were used. CTG is still the most widely employed prenatal technique for monitoring fetal health.

Artificial Intelligence has made rapid pace in the previous decades, and it has been successfully applied in several important fields. The worldwide demand of ML techniques is used in every domain. By using the datasets and inaccurate medical records, the analysis of the new cases for the development of decision support system will run into missing data which are the common executions. ML allows diagnostic systems to be faster and more precise than humans.

2   Literature Survey

In this paper, fetal health prediction based on machine learning approach, we used clinical datasets of the fetus, to predict the fetal status if the condition is normal, abnormal or chronic. According to the survey done, we have come through various techniques for determining the fetal health.

In [1], the author has used the methodology the CNN and long short-term memory to determine the image-based motion tracking method based on deep learning to identify the fetal motion and movements directly from the obtained images.

In [2], they have used a supervised method for estimation of acoustic shadow region in fetus by using the ultrasound imaging. They have integrated the shadow confidence maps by classifying the multi-view image fusion. Novel-based CNN is used to determine the pixel-wise 2D images. Diagnostic methods are used based on the transvaginal ultrasound where the length of the cervix is examined; also, K-nearest neighbors (KNN) and convolutional neural networks were used to predict the preterm birth [3]. The fetal weight prediction is being done based on using ultrasound image that is using fuzzy C means clustering; also, the weight can be calculated based on abdomen circumference and biparietal diameter [4]. Spatio-Temporal Neural Networks were used to find fetal health and stress. There is unusual appearance in different anatomical structures during various phases of heart cycle [5]. Fetal movement counting provides valuable information where a low complexity medium for fetal movement detection is being shown based on various factors [6]. A convolutional neural network is used to predict a baby's risk of oxygen deprivation during the childbirth. They have used a machine learning technique to achieve the

data. As a result, they comprise the clinical data for future enhancement [7–17].The fetal growth state is preclassified based on developing a predictive classifier model using machine learning algorithms, and accuracy is improved using sonography method to ultrasound examination, up to 77% [18]. The above-stated techniques are not able to predict the accurate fetal growth.

To overcome this existing system, use classification algorithms in ML to predict the accuracy as normal, abnormal or chronic. By applying the conventional method, AI and ML approaches and image classification methods, designing an accurate model for the early detection of abnormalities is important for the fetal well-being.

3 Proposed System and Methodology

The standard AI and ML method for measuring fetal health and fetal movement is based on a non-invasive stress test that includes a No. of fetal movements over a period of time. Fetal MRI is also one of the most challenging and emerging applications for uncontrolled and abnormal baby movements. The maximum accuracy of the guess can be achieved by collecting a sufficient amount of dataset with over 2126 records for both normal and unusual attribute reading data. For standard sonographer's ultrasound, testing depends on the appearance of the defined anatomical structures. This is achieved by SVM in AI and ML. In ML, data collection methods can be divided into basic data collection methods (1) and secondary data collection methods (2). In (1), it is collected directly and has never been used in the past. Using the (1), data collected is accurate and clear and has been used in the past. Researchers can obtain data from sources within and outside the organization. This paper focuses on using secondary data acquisition. It consists of data preparation, data processing, feature engineering and prediction algorithms for finding the accuracy of results. The overall process is mentioned in Fig. 1.

Data Preparation
Fetal health dataset is used in good deal of research. Due to the insufficient information of fetal, the sole method to manage the model and to put together the acquired data from a reliable source. As a result, the fetal dataset was worn to generate the dataset. This dataset is used to predict fetal health by utilizing different features of the fetus. For ML and Data Visualization applications, to settle on a subset of the initial data the filtering method is used.

Data Pre-processing
A dataset is formed from various paradigm. A set of features differentiate a data item, understanding the basic features of an item. To refine the dataset quality, the starting step would be to get rid of the empty values which examine a drag while fetching the accuracy. Furthermore, the outline of the dataset is necessary to consider and to urge a start for the feature engineering process.

Fig. 1 Dataset processing
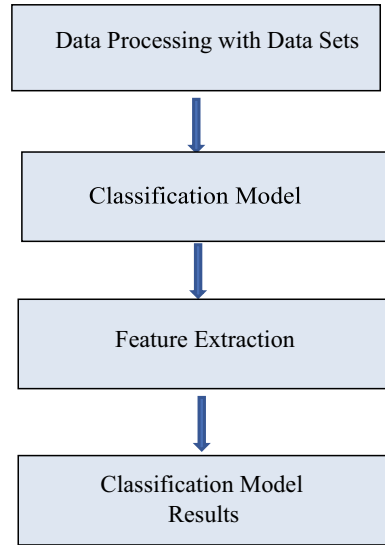
Feature Engineering

Feature engineering performs a key role in the methodology. Firstly, visualization conveys about the various health parameters and obtains the understanding whether it is an essential attribute or not. Then, the model is employed w.r.t every attribute with itself and also with the contrasting features.

Prediction Algorithms

It is based on (above A, B and C) to estimate the fetal health status, whether it is normal or pathological where maternal clinical data was questioned. Classification algorithms are used. The high prediction accuracy can be achieved by collection of adequate number of datasets by over 100,000 observations of both normal and abnormal data for attribute learning. To build accuracy of every classifier, the algorithm which has high accuracy is employed for prediction. Table 1 represents the development environment ML-based Conventional Classifier. The following components are used to predict the fetal health condition.

Table 1 Development environment

Components	Description
System	Windows 10 Pro 64 bit
Server	CherryPy WSGI Server
IDE	Jupyter Notebook
Browser	Chrome, Firefox, IE
Library and framework	CherryPy

4 Results and Discussion

After applied the ML techniques for various datasets of fetal heart rate, acceleration, fetal movement, uterine contractions and the overall fetal health is monitored. Figure 2 shows the baseline fetal heart rate, in which the range 110–160 beats per minute is considered as 'normal' and baseline less than 110 bpm is 'abnormal' or 'slow heart rate'.

Figure 3 shows the fetal heart rate acceleration per second. The short-term increase in the heart rate of at least 15 beats per min or lasting at least 15 s in 'x-axis' is considered as 'normal', and other cases are considered as 'abnormal'. It can obtain by checking on the oxygen supply w.r.t to mm Hg in 'y-axis'. Figure 4 showcases the number of fetal movements per second where 10 movements in an hour or less are 'normal'. If the movement is not recorded at the stipulated time, then it is considered as 'abnormal'. In Fig. 5, uterine contraction for the 'normal' condition is 0.005 ms. For the suspect, it is 0.001 ms and 0.003 ms for 'pathological' or 'abnormal'. Figure 6 shows the overall effect of the parameters on the fetal health condition where it shows 1 as 'normal' condition w.r.t to heart rate, where above 1 is considered as 'suspect' or 'pathological'.

Based on the input of the parameters, our technique predicts 92–98% as normal and 2–4% as suspect or pathological as shown in Fig. 7.

The comparative study of our technique with existing platform is shown in Table 2. [1] Utilized CNN technique, shows fetal movement is less than 15 beats/min, fetal heart rate is normal and overall prediction rate is < 90%, when compared with our proposed mechanism in which fetal movement is < 15 beats/min, heart rate is normal and overall predictions are 92–98% and other existing technique shows less significant results.

Fig. 2 Baseline value
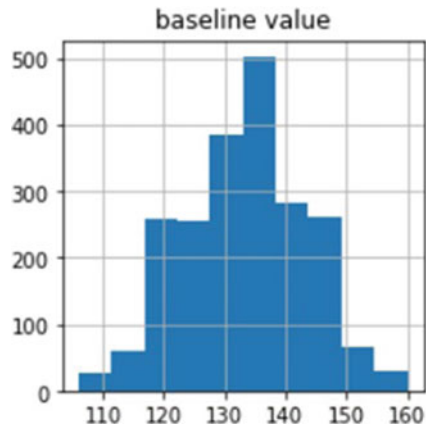
Fig. 3 Fetal heart rate
acceleration
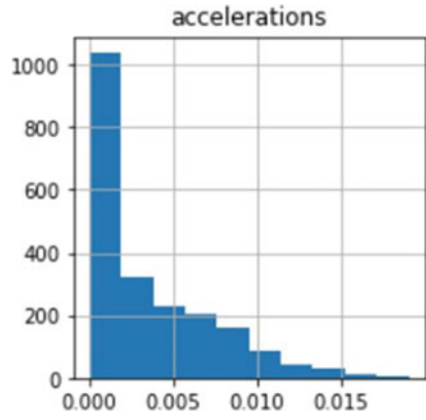
Fig. 4 Fetal movements
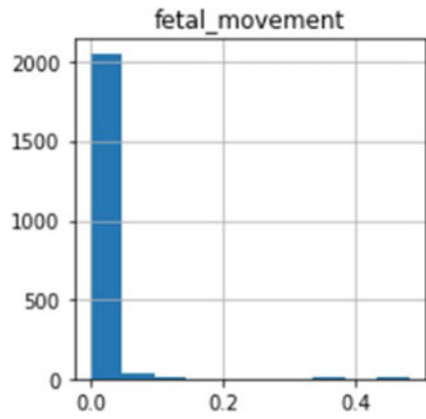
Fig. 5 Uterine contraction

Fig. 6 Fetal health
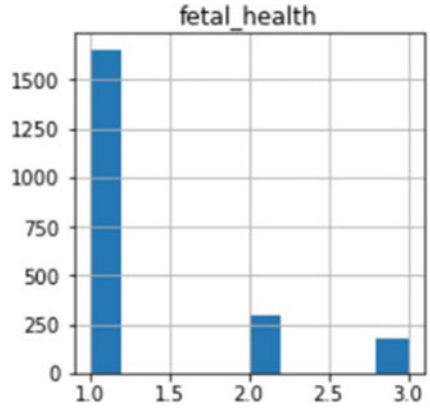
Fig. 7 Fetal health with results
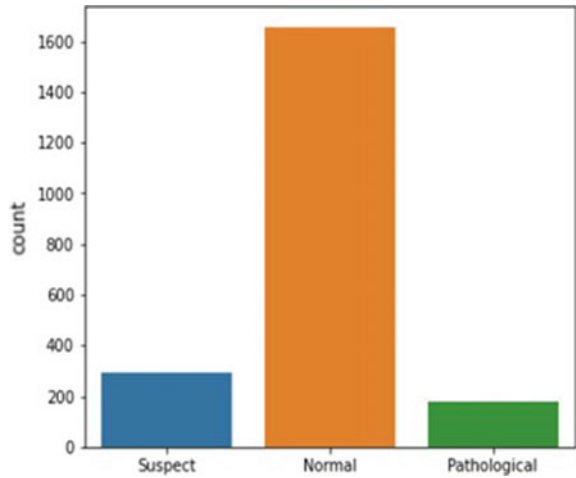
Table 2 Comparative study of our technique with existing techniques

Name	Techniques	Fetal movement (beats/min)	Fetal heart rate	Overall prediction (%)
[1]	CNN	<15	Normal	<90
[2]	Novel-based CNN	>15	Slower heart rate	60
[4]	CNN predictive classifier	>15	Slower heart rate	70
Proposed	ML with conventional algorithm	<15	Normal	92–98

5   Conclusion

A major contributor to this work is the e-health support program for nurses and pregnant patients that helps to predict birth defects associated with conventional machine learning methods. We analyzed various mechanisms in the survey. Based on the survey, we proposed ML-based conventional technique to predict the fetal growth and its condition. Our results proved 92–98% of normal fetal growth. With the help of this mechanism, the services to pregnant women and therapist or physicians can evaluate each patient's case in detail, based on the individual patient's limitations. We also propose guidelines for future research into child health assessment using machine learning strategies.

References

1. Singh A (2020) Deep predictive motion tracking in magnetic resonance imaging: application to fetal imaging. IEEE Trans Med Imaging 39(11):3523–3534
2. Meng Q, Hou (2019) Weakly supervised estimation of shadow confidence maps in fetal ultrasound imaging. IEEE Trans Med Imaging 38(12):2755–2767
3. Rakesh Raj (2021) A Machine learning methods for preterm birth prediction:Hindawi J Healthc Eng 1–11
4. Pradipta GA (2017) Fetal weight prediction based on ultrasound image using fuzzy C means clustering and iterative random Hough transform
5. Lee LH, Alison Noble J (2020) Automatic determination of the fetal cardiac cycle in ultrasound using spatio-temporal neural networks. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI)
6. Rooijakkers MJ (2016) Feasibility study of a new method for low-complexity fetal movement detection from abdominal ECG recordings. IEEE J Biomed Health Inform 20(5):1361–1368
7. Christopher WG (2019) Multimodal convolutional neural networks to detect fetal compromise during labor and delivery. IEEE Access 7:112026–112036
8. Deressa TD, Kadam K (2018) Prediction of fetal health state during pregnancy: a survey. Int J Comput Sci Trends Technol (IJCST) 6(1)
9. Marques JAL (2021) IoT-based smart health system for ambulatory maternal and fetal monitoring. IEEE Internet Things J 8(23):16814–16824
10. Qu R, Xu G (2020) Standard plane identification in fetal brain ultrasound scans using a differential convolutional neural network. IEEE Access 8:83821–83830
11. De Jonckheere J, Garabedian C, Charlier P, Champion C, Servan-Schreiber E, Storme L, Debarge V, Jeanne M, Logier R (2017) Influence of ECG sampling rate in fetal heart rate variability analysis. In: Annual international conference of the IEEE Engineering in Medicine and Biology Society
12. Monkaresi H, Calvo RA, Yan H (2013) A machine learning approach to improve contactless heart rate monitoring using a webcam. IEEE J 18(4):1153–1160
13. Cömert Z (2017) Comparison of machine learning techniques for fetal heart rate classification. Acta Phys Pol A 132(3):451–454
14. Biswas D, Everson L (2019) CorNET: deep learning framework for PPG-based heart rate estimation and biometric identification in ambulant environment. IEEE Trans 13(2):282–291
15. Hsu G-SJ (2020) A deep learning framework for heart rate estimation from facial videos. Neurocomputing 417:155–166
16. Maragatham G, Devi S (2019) LSTM model for prediction of heart failure in big data. JMS 43(5):111

17. Bian M, Peng B (2019) An accurate LSTM based video heart rate estimation method. In: PRCV. Springer, Xi'an, China, Nov 2019, pp 409–417
18. Chinnaiyan R, Alex S (2021) Machine learning approaches for early diagnosis and prediction of fetal abnormalities. ICCC

BLDC Motor and Its Speed Characteristics Analysis Based on Total Harmonic Distortion

K. M. N. Chaitanya Kumar Reddy, N. Kanagasabai, and N. Gireesh

Abstract The work brings up the system modeling of a Brushless DC motor and later the study on its responses using the conventional controller like PID with the fuzzy PID controller. The system's speed characteristics were studied with respect to change in load. Based on the parameters, the system's response to reach the steady state is analyzed using MATLAB simulation tool. The total harmonic distortion developed in the Brushless DC motor is calculated at different conditions like load, no load, and reduced load using the frequency analyzer. The results were evaluated by implementing the above said controllers, and the comparison was done in the conclusion.

Keywords Brushless DC motor · THD · Steady state · Fuzzy

1 Introduction

In current situation, motors play a vital role in many applications. Brushless DC motors are the commutated motors which work on a DC supply. In the applications of vehicles manufacturing and in the fields of robotics, BLDC motors are prominently used. During the operation of a BLDC motor, the problem arises because of vibrations which we call it as Total Harmonic Distortion (THD). This may be due to effect of change in reference speed, variations in load, and other parameters. Hence for minimization of these disturbances, we apply different controllers during the modeling process, and hence, we get the steady state responses.

K. M. N. C. K. Reddy (✉) · N. Kanagasabai
Annamalai University, Chidambaram, India
e-mail: chaitanyakmn@vidyanikehan.edu

N. Gireesh
Sree Vidyanikethan Engineering College, Tirupati, India

2 Literature Survey

From the papers referred [1, 2], I have done the system modeling of Brushless DC motor. The parameters were chosen based on fixed values studied during the survey from the initial three papers mentioned in the references [1–3]. The parameters are considered as inertia, resistance, and load. Based on tuning process referred from the references [4–6], the kp, ki, and kd values were considered under the parameters of PID controller. The fuzzy rule base was developed based on the input–output values relations studied from the [7, 8]. The simulation was developed from the practice using the MATLAB software [9–11]. The literature survey says that the fuzzy controller-based system shows better response when compared with the system using PID controller. The total harmonic distortion information was collected from the Internet and based on its characteristics the responses were collected from the articles presented in the references. In the paper, the same comparison was done with the developed BLDC system, and the comparison was made using a graph.

3 Block Diagram

Figure 1 describes the block diagram of BLDC servomotor drive system. The IGBT inverter is provided with a DC power supply, and the output from the inverter would be an AC supply. This supply will be provided to BLDC motor. The fuzzy PID and PID controllers are provided with certain gains to which it is evaluated under various environments of load and speed alterations. The parameters include resistor, inductor, and capacitor values of the BLDC system. Along with the values, the other parameters include moment of inertial and the load conditions. Based on these parameters, the corresponding system responses were observed for the considered controllers using soft computing.
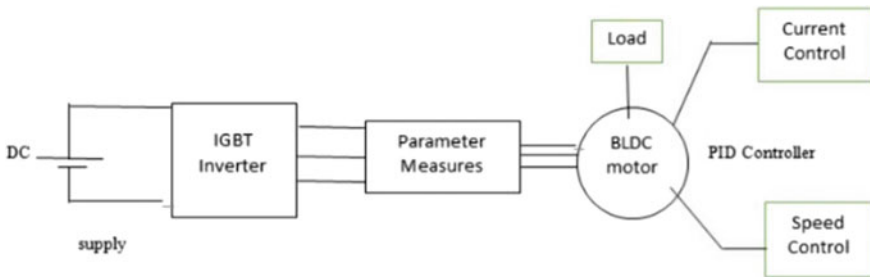
Fig. 1 BLDC servomotor drive system

4 Brushless DC Servomotor Drive System Modeling

The IGBT inverter is provided with a DC power supply and the output from the inverter would be an AC supply. This supply will be provided to BLDC motor. The fuzzy PID and PID controllers are provided with certain gains to which it is evaluated under various environments of load and speed alterations. From the equivalent circuit of BLDC motor shown in Fig. 2, the voltages between the lines in the equivalent circuit of a BLDC are mathematically given by

$$
\begin{bmatrix} V_{ab} \\ V_{bc} \\ V_{ca} \end{bmatrix} = \begin{bmatrix} Ra & -Ra & 0 \\ 0 & Rb & -Rb \\ -Rc & 0 & -Rc \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}
$$
$$
+ \begin{bmatrix} La - Mb & Mb - La & 0 \\ 0 & Lb - Mc & Mb - Lc \\ Mc - La & 0 & Lc - Ma \end{bmatrix}
$$
$$
\times \frac{di}{dt} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} e_a - e_b \\ e_b - e_c \\ e_c - e_a \end{bmatrix} \tag{1}
$$

where L and M are the self and mutual inductances of each line a, b, and c, respectively. R is the winding resistor. The electromotive forces are e_a, e_b and e_c; i_a, i_b and i_c are the phase streams of individual lines.

As the self-inductance is given high priority when compared with the mutual inductance, the same matrix equation can be evaluated as
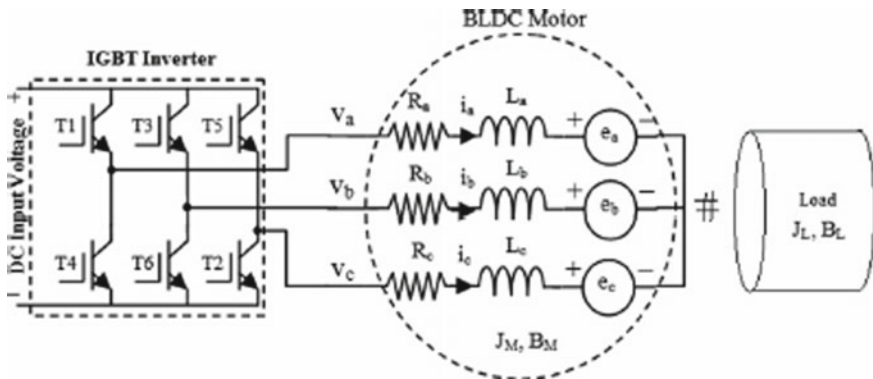
Fig. 2 Equivalent circuit

$$\begin{bmatrix} V_{ab} \\ V_{bc} \\ V_{ca} \end{bmatrix} = \begin{bmatrix} Ra & -Ra & 0 \\ 0 & Rb & -Rb \\ -Rc & 0 & -Rc \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}$$

$$+ \begin{bmatrix} La & -La & 0 \\ 0 & Lb & -Lb \\ -Lc & 0 & Lc \end{bmatrix}$$

$$\times \frac{di}{dt} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} e_a - e_b \\ e_b - e_c \\ e_c - e_a \end{bmatrix} \tag{2}$$

The torque developed is

$$T = (e_a i_a + e_b i_b + e_c i_c)/\omega \tag{3}$$

where $\omega = K_t I$ and $I = i_a = i_b = i_c$, ω is the angular velocity and torque constant is symbolized as K_t. As the electromagnetic torque can be applicable to overwhelm the conflicting load and inertia torques.

$$T_e = T_L + J_M \frac{dw}{dt} + B_L \omega \tag{4}$$

Here, load torque factor is indicated by T_L, inertia factor is indicated by J_M, and friction factor of the BLDC servomotor is B_L. The T_L, is communicated by J_L and B_L components as

$$T_L = J_L \frac{dw}{dt} + B_L \omega \tag{5}$$

The yield power will be

$$P = T_e \, \omega \tag{6}$$

5   Design of Controllers

The values of P, I and D were considered and its gain values were evaluated as $K_p = 11$, $K_i = 5$, and $K_d = 0.1$ based on the references. The tuning method done was based on Zieglar–Nicholas tuning for the PID controller.

The considered second controller is a combination of PID and fuzzy logic with Sugeno rule base. It is basically a combination developed through rule base and inference under uncertain conditions.
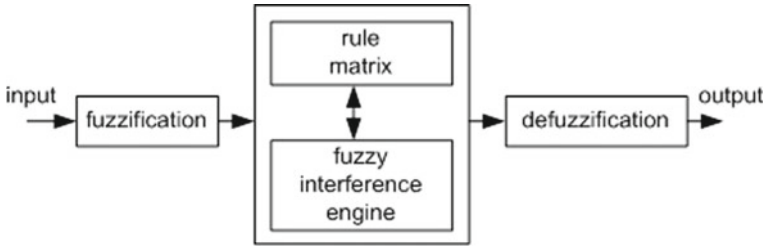
Fig. 3 Fuzzy logic controller

The rule base developed under uncertainty conditions is normally evaluated with the human thinking levels, and the developed systems under these conditions showcase nonlinearity in handling the various industries. The fuzzy-based controller is normally developed using fuzzification, rule matrix, inference system, and defuzzification which constitutes the controller. The block diagram of a fuzzy logic controller is shown in Fig. 3.

Major stage of the process involved in fuzzy takes place in contributing the relation between input factors under considered fuzzy system and the output result. The considered factors stand fuzzified by applying pre-defined input membership functions. All these membership functions may vary in shapes. The utmost applied includes triangular, trapezoidal, sinusoidal, and exponential shape membership functions. Modest functions won't need intricate calculating besides won't excess the operation. The grade of membership function is resulted through placement of selected input variable on the parallel axis, whereas perpendicular axis indicates mark of membership of the input variable. A mere condition is that membership function need to run into range of zero and one. Zero represents that input variable doesn't belong to fuzzy set, whereas the value one means fully a member of the fuzzy set.

With each input parameter, there will be a unique membership function connected. These functions possess a weighting factor by values of input and the operative rules, respectively. The degree of membership of each active rule is evaluated by these weighting factors. The surface view of fuzzy rule base is shown in Fig. 4, and membership functions considered are as shown in Figs. 5 and 6. The rule editor is shown in Fig. 7.

6 Simulation Results

The BLDC motor is applied with constant speed, varying speed, and load input conditions. The corresponding responses of the Brushless DC motor and control system parameters are measured and analyzed. The developed simulation circuit of BLDC motor is shown in Fig. 8.
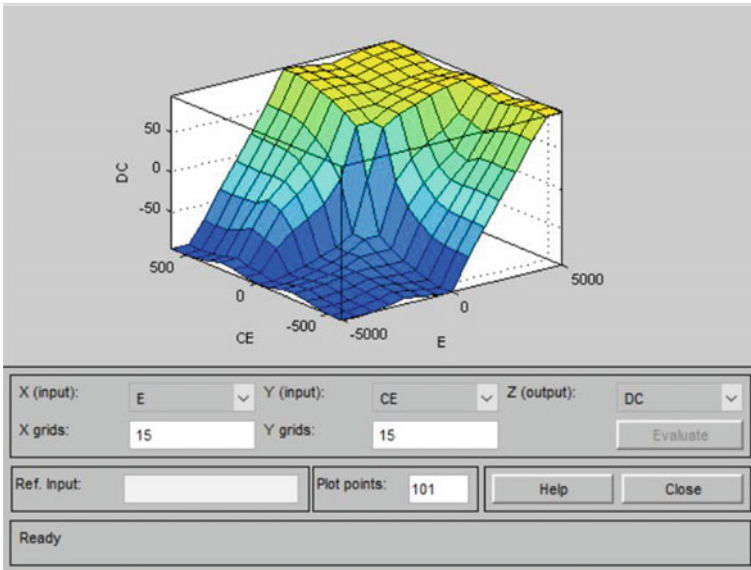
Fig. 4 Surface view of fuzzy rule base
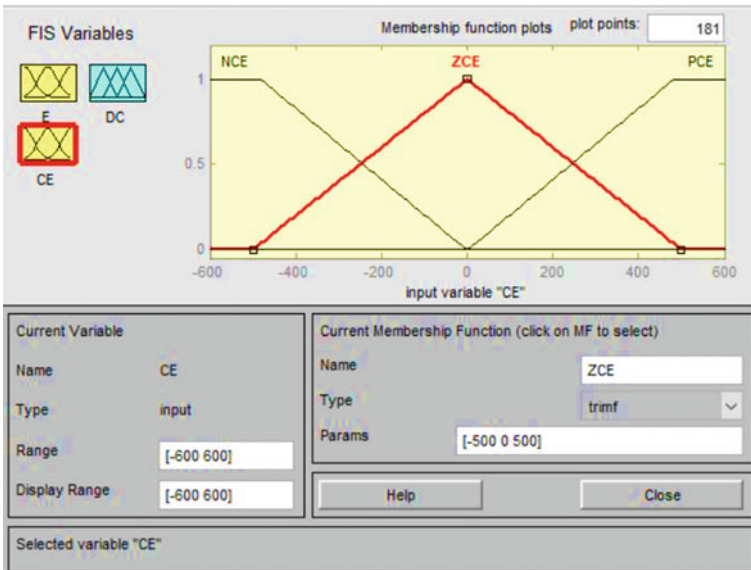
Fig. 5 Change in error as one input membership function
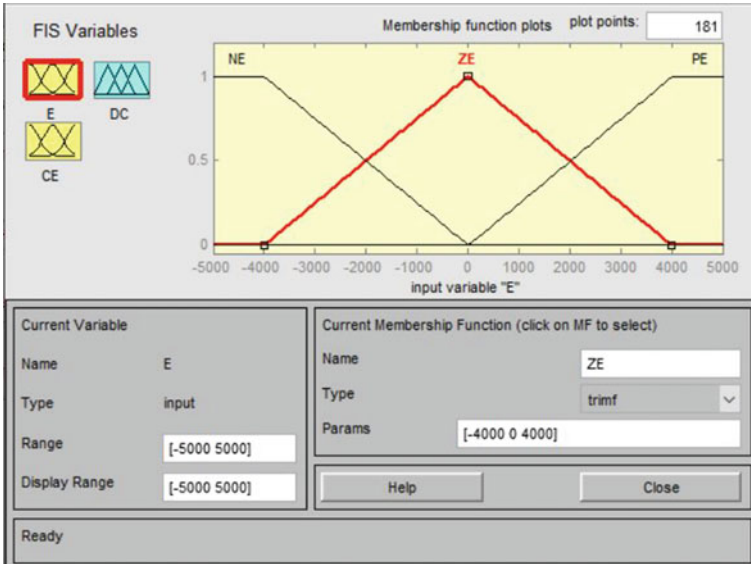
Fig. 6 Error as other input membership function
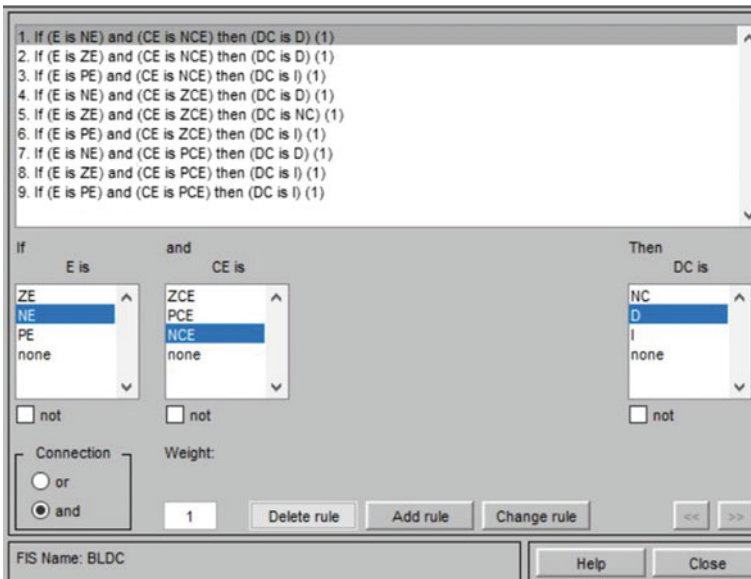
Fig. 7 Rule editor
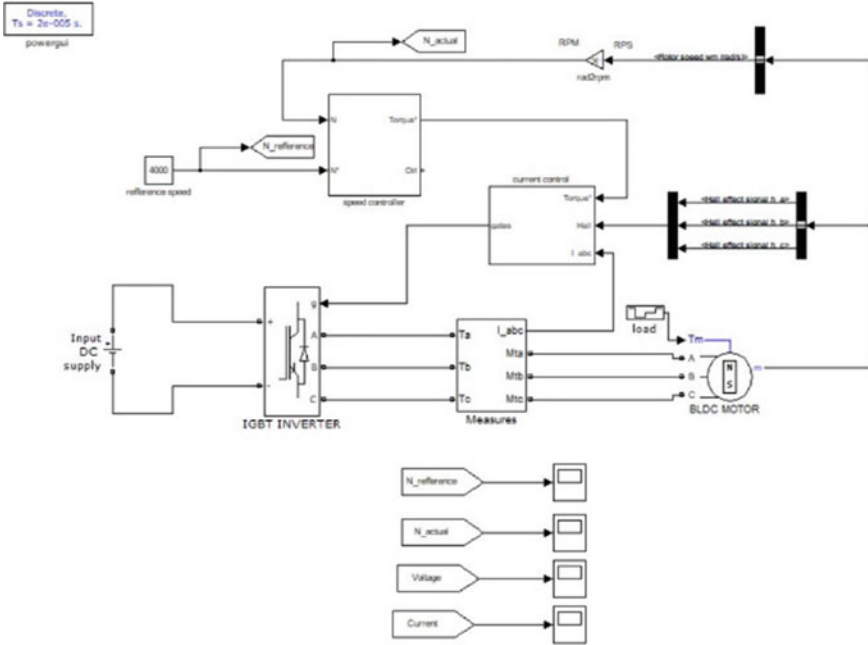
Fig. 8 Simulation diagram developed for the BLDC motor

The considered specifications of the system are given in Table 1. The responses of the system based on THD harmonics using PID and Fuzzy PID controllers are shown in Figs. 9, 10, 11, and 12. The responses were calculated based on the different configurations of parameters like load 'R' and inertia 'J'.

Table 1 System specifications	Rated voltage	36 V
	Rated current	5A
	No. of poles	4
	No. of phases	3
	Rated speed	4000 RPM
	Rated torque	0.42 N.m
	Torque constant	0.082 N.m/A
	Mass	1.25 kg
	Inertia	$23e{-}06$ kg-m^2
	Resistance per phase	0.57 ohms
	Inductance per phase	1.5 mH
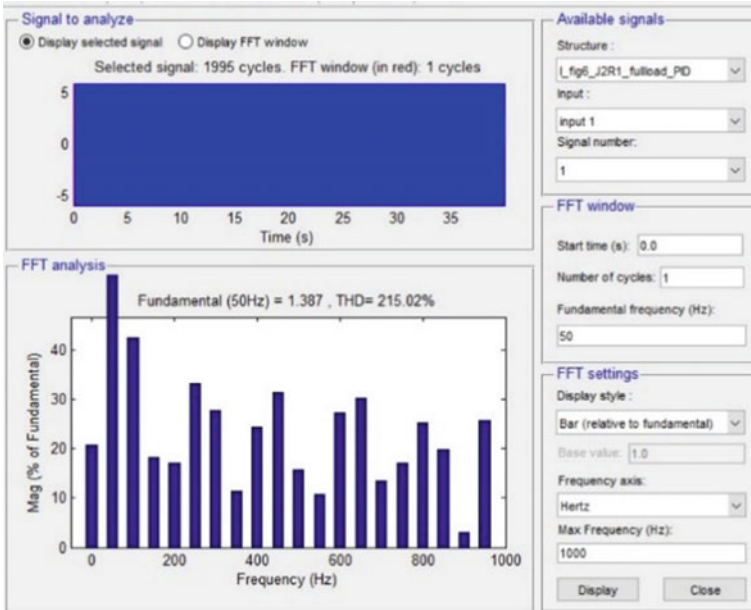
Fig. 9 PID controller response of J2 and R1 with full load

Fig. 10 Fuzzy PID controller response of J2 and R1 with full load
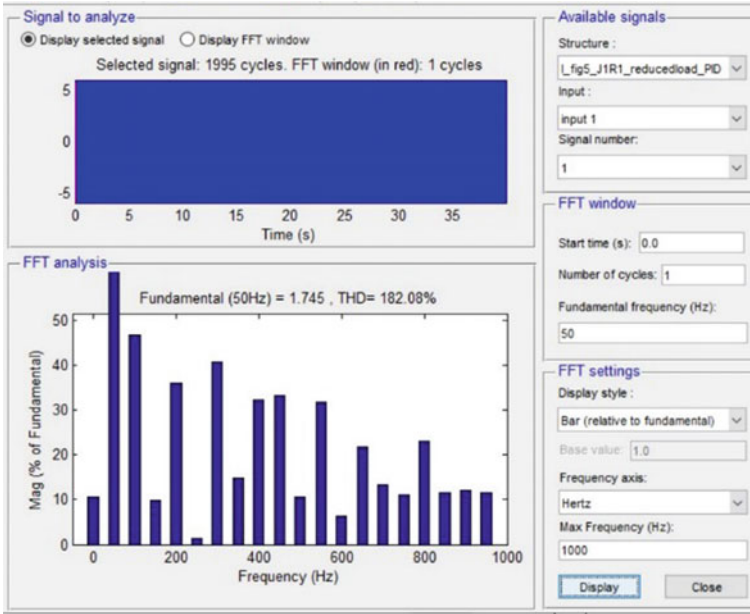
Fig. 11 PID controller response of J1 and R1 with reduced load
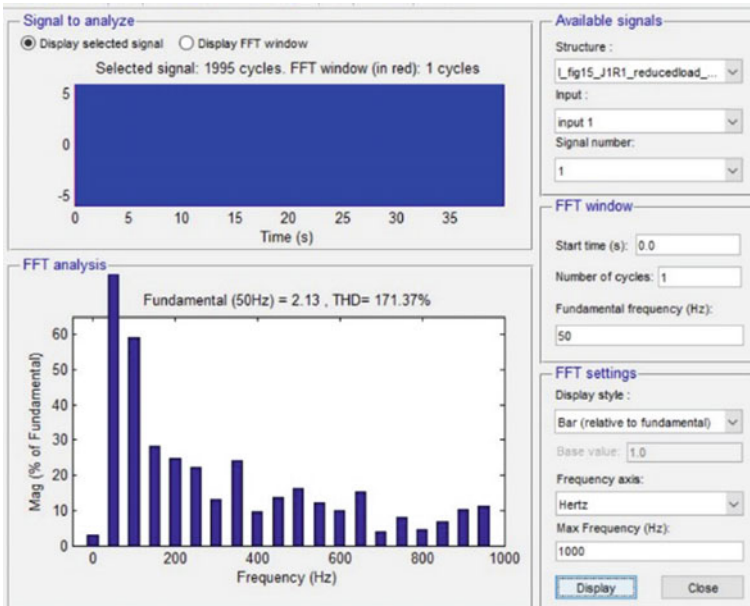
Fig. 12 Fuzzy PID controller response of J1 and R1 with reduced load

Fig. 13 Comparison of THD through PID and fuzzy PID
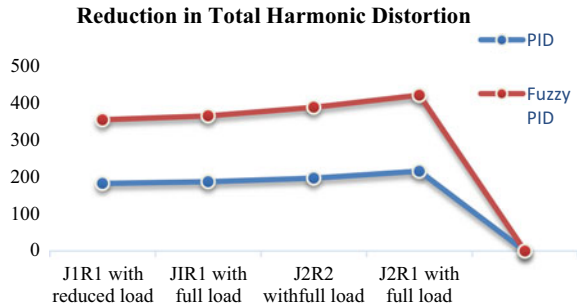
7 Conclusion

In this work, the designed BLDC system was simulated using the PID and fuzzy controllers in MATLAB. From the responses, it can be observed that the performance of fuzzy controller seems to slightly better when compared to PID controllers under parameter variations, resulting in reduction of the harmonic distortion which can be observed from Fig. 13.

References

1. Shanmugasundram, Zakariah, Yadaiah (2009) Low-cost high performance brushless dc motor drive for speed control applications. In: 2009 International conference on advances in recent technologies in communication and computing, Kottayam, India, Oct 27–28, pp 456–460
2. Shanmugasundram, Zakariah, Yadaiah (2009) Digital implementation of fuzzy logic controller for wide range speed control of BL DC motor. In: International conference on vehicular electronics and safety (ICVES), Pune, India, Nov 10–12, pp 119–124
3. Raju SS et al (2013) Implementation of PID and Fuzzy PID controllers for temperature control in CSTR. Int J Adv Res Comput Sci 4(5):12–17
4. Reddy KMNCK, Rajani B, Raju PS Control of non linear two mass drive system using ANFIS. Tc 2:1
5. Horvat R, Jezernik K (2014) An event-driven approach to the current control of a BLDC motor based FPGA. IEEE Trans Ind Electron 61(7):3719–3726
6. Kandiban R, Arulmozhiyal R (2012) Speed control of BLDC motor using adaptive fuzzy PID controller. Procedia Eng 38:306–313
7. Kim I et al (2010) Compensation of torque ripple in high performance BLDC motor drives. Control Eng Pract 18(10):1166–1172
8. Yigit T, Celik H (2020) Speed controlling of the PEM fuel cell powered BLDC motor with FOPI optimized by MSA. Int J Hydrogen Energy 45(60):35097–35107
9. Im H, Yoo HH, Chung J (2011) Dynamic analysis of a BLDC motor with mechanical and electromagnetic interaction due to air gap variation. J Sound Vib 330(8):1680–1691
10. Yamashita RY et al (2018) Comparison between two models of BLDC motor, simulation and data acquisition. J Braz Soc Mech Sci Eng 40(2):1–11
11. Priyanka CP, Jagadanand G (2021) Design and analysis of BLDC motor for electric vehicle application. In: Advances in automation, signal processing, instrumentation, and control. Springer, Singapore, pp 977–985

12. Gireesh N, Sreenivasulu G (2014) Comparison of PI controller performances for a conical tank process using different tuning methods. In: 2014 International conference on advances in electrical engineering (ICAEE), pp 1–4. https://doi.org/10.1109/ICAEE.2014.6838426
13. Vadivazhagi S, Jaya N Fuzzy gain scheduled PI controller for a two tank conical interacting level system. Int J Eng Technol (IJET) 6(6):2588–2594
14. Tang K-S et al (2001) An optimal fuzzy PID controller. IEEE Trans Ind Electron 48(4):757–765

Calculating the Traffic Density of Real-Time Video Using Moving Object Detection

S. Rakesh and Nagaratna P. Hegde

Abstract Moving object detection (MOD) is one of the more prominent fields of research now a day because the application areas of it increasing rapidly. Several applications of MOD are recognizing and tracing vehicles or animals or human beings, finding the speed of objects or vehicles etc. Traffic controlling is one of the prodigious issues facing by most of the great cities in the world especially in India all metropolitan cities suffering with this issue. Therefore, there is a huge demand to introduce an intelligent traffic controlling system which works by finding the live traffic and changing the traffic lights by its own intelligence. In this paper, proposing a paradigm to find the live traffic density using MOD algorithm which interns uses prominent background subtraction practice by giving a traffic video as input to the proposed model so that this can work efficiently with live traffic video.

Keywords Traffic video · Traffic density · Moving object detection · Background subtraction

1 Introduction

At present the world has plenty of digital ocular information. Several image analysis techniques are used to understand and analyzing this extensive collection of information especially image data. One of the very important paradigms is object detection, which has many real-time applications. Finding the theft vehicles, finding the criminals, video surveillance systems, and detecting and tracking objects are some of the applications of object detection paradigm. There are several techniques for detecting the objects present in an image. The different types of object detection methods, i.e.,

S. Rakesh (✉)
Chaitanya Bharathi Institute of Technology, OU, Hyderabad, India
e-mail: srakesh_it@cbit.ac.in

N. P. Hegde
Vasavi College of Engineering, Hyderabad, India
e-mail: nagaratnaph@staff.vce.ac.in

959
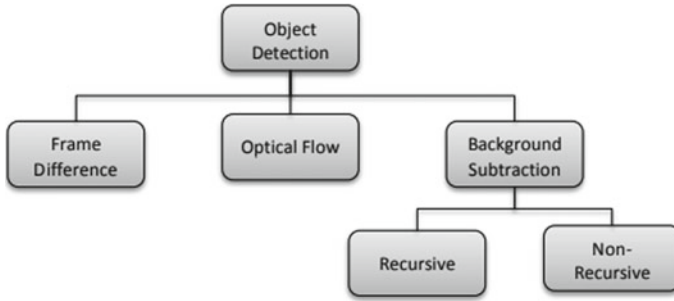
Fig. 1 Types of object detection techniques

frame difference method, optical flow technique, recursive, and non-recursive background subtraction methods are shown in Fig. 1. The probable usefulness of images can be acquired by using these object detection methods. The principal information of image is the objects present in it. Therefore, best object detection [1] techniques are very much needed for many real time applications.

Moving objects detection (MOD) is a big challenge. In this paper, proposing a method to find the moving vehicles near the traffic light signal using MOD method, which can be helpful in controlling the traffic lights dynamically based on the number of vehicles detected or based on the moving vehicles density at the traffic light signal.

2 Literature Review

Moving objects can be detected using several techniques. Some of the techniques used to detect the moving objects are recursive and some are non-recursive techniques. Mixture of Gaussians (MoG), Kalman filter, and approximation media filter are examples of recursive moving object methods. Linear predictive filter, frame differencing, and median filtering are non-recursive moving object detection techniques. These recursive and non-recursive techniques briefly elaborated by Chandrasekhar et al. [2]. The MoG algorithm for detecting foreground and also few important advancements to overcome some critical circumstances were explained by Yilmaz et al. [3] and Bouwmans et al. [4]. Live video surveillance using background modeling is presented in Jeeva and Sivabalakrishnan [5]. The traditional approaches and latest techniques of background modeling are summarized in Bouwmans [6]. Lim et al. [7] presented the moving object detection especially human movement recognition with fuzzy logic.

One more very important real-time application of moving object detection paradigm is smart traffic control and management system, i.e., controlling the traffic lights by calculating live moving vehicles near the traffic light junction. Some of the moving vehicles recognition techniques are optical flow estimation method presented in [8, 9], subtracting background paradigm is discussed in [10–12] and the frame

differencing method presented in [13–15]. Out of these three techniques, the best and mostly used method for intelligent traffic controlling system is background subtraction paradigm. Optical flow method is not suitable for traffic domain because it requires more complicated computation. Frame differencing technique is also not used in traffic domain; the reason is this technique is more sensitive with illumination changes which creates difficulties in traffic domain.

3 Methodology

In this paper, the datasets used are the traffic videos downloaded from kaggle. The key steps of proposed methodology are identifying region of interest (ROI), generating frames from the input video, apply createbackgroundsubtractormog2 () function for detecting moving vehicles, performing masking to highlight the detected moving vehicles, drawing contour lines, i.e., rectangles over detected vehicles and then finally finding the area occupied by these rectangles, i.e., vehicle density in the frame. In this process, the intermediate results are saved and observed for performance analysis. The ROI highlighted images for the video1 and video 2 datasets are shown in Figs. 2 and 3, respectively.

Fig. 2 ROI for video1 dataset

Fig. 3 ROI for video2 dataset
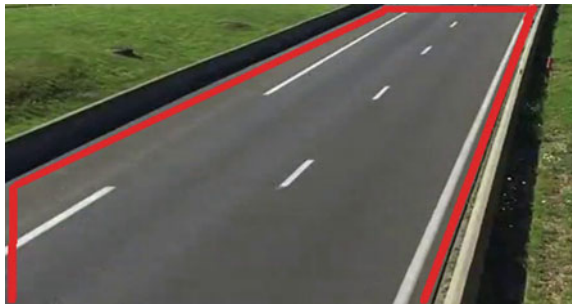
The result images of applying background subtraction and masking are shown in Figs. 5 and 7. These images are the results of respective input frames shown in Figs. 4 and 6. Moving objects, i.e., vehicles are highlighted in these images with white pixels. And, then contour lines drawn sample results are shown in Figs. 9 and 11. These results are the outputs of the masked frames shown in Figs. 8 and 10, respectively.

Fig. 4 Sample frame from traffic video

Fig. 5 Vehicle detected image for the sample frame in Fig. 2
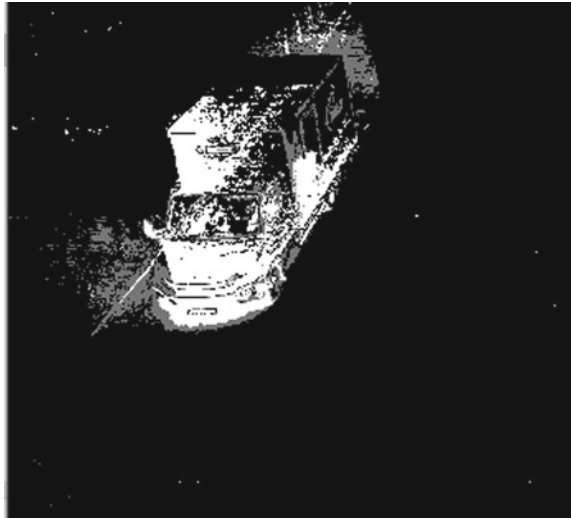
Fig. 6 Sample frame from traffic video

Fig. 7 Vehicles detected image for the sample frame in Fig. 4
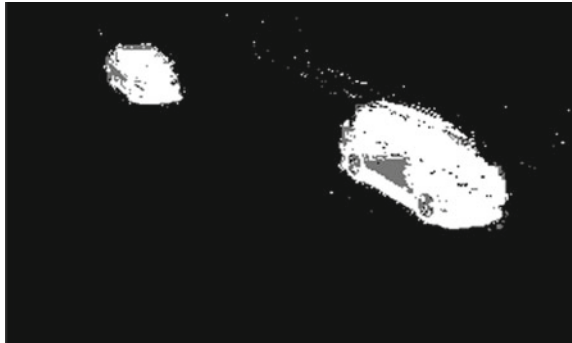
Fig. 8 Vehicles detected image

Fig. 9 Contour lines drawn over detected vehicles for Fig. 8
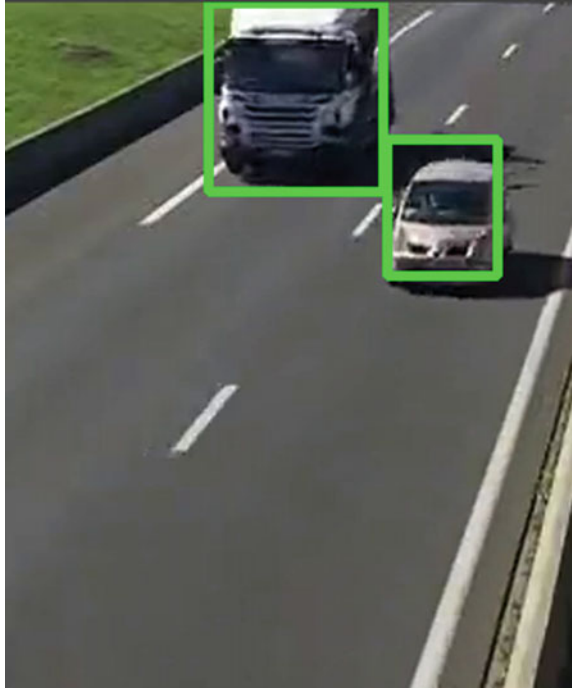
Fig. 10 Vehicles detected image

Fig. 11 Contour lines drawn over detected vehicles for Fig. 10
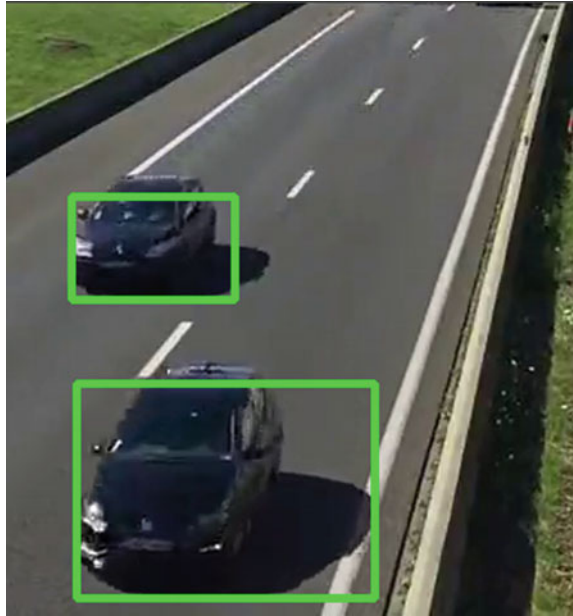
4 Experimental Results

The proposed methodology working is tested by processing three video datasets through proposed model. The sample results with traffic densities values displayed on the live images generated from videos are shown in Figs. 12, 13, and 14. Figure 12 displays the traffic density value 31, the meaning of it is vehicles occupied 31% of ROI area. Similarly, the Fig. 13 describes 29% of ROI area occupied by the vehicles shown in the image. According to manual validation the results are almost maintaining a very good accuracy. Figure 14 displaying the traffic density value as 0 when there are no vehicles, this is absolutely correct and can be considered as an extreme test case. Proposed model is considering shadows also as moving vehicles, i.e., shown in Fig. 11, in such cases the traffic density values are more than the original values. This is one drawback of proposed model; this can be overcome by using suitable latest technologies. Apart from it this model works efficiently in real time.

5 Conclusion and Future Scope

The proposed model in this paper gave good results which are mentioned in results section. By using this proposed method, calculated the density of the live traffic by giving traffic videos as input. Actually, in this paper, generated 30 fps and each frame density calculated within a second time. The average traffic density of latest

Fig. 12 Sample result with traffic density value 31
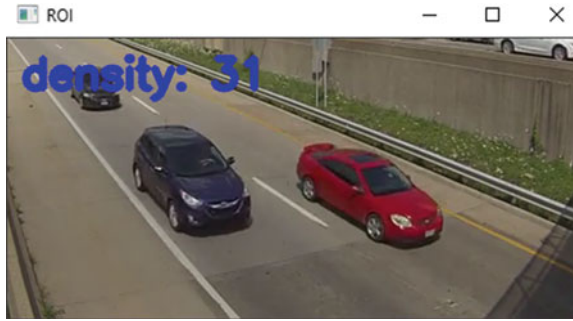
Fig. 13 Sample result with traffic density value 29

Fig. 14 Sample result with traffic density value 0

few seconds needs to be calculated so that which can be used to control the traffic lights dynamically, i.e., if this value is less than some fixed threshold value then traffic light signal need to change to red because the value indicates less traffic else

green traffic light need to be continued until reach max allocated time for example 60 s. After 60 s or max allotted time reached by default traffic light changes its signal. In this way the proposed system may solve the issues related to traffic congestion problems. The future need of this proposed model is to find the traffic density more accurately by removing the shadows of the vehicles.

References

1. Tiwari M, Singhai R (2017) A review of detection and tracking of object from image and video sequences. Int J Comput Intell Res 13(5):745–765
2. Chandrasekhar U, Das T (2011) A survey of techniques for background subtraction and traffic analysis on surveillance video
3. Yilmaz A, Javed O, Shah M (2006) Object tracking: a survey. ACM Comput Surv (CSUR) 38(4):13
4. Bouwmans T, El Baf F, Vachon B (2008) Background modeling using mixture of Gaussians for foreground detection—a survey. Recent Pat Comput Sci 1(3):219–237
5. Jeeva S, Sivabalakrishnan M (2015) Survey on background modeling and foreground detection for real time video surveillance. Procedia Comput Sci 50:566–571
6. Bouwmans T (2014) Traditional and recent approaches in background modeling for foreground detection: an overview. Comput Sci Rev 11:31–66
7. Lim CH, Vats E, Chan CS (2015) Fuzzy human motion analysis: a review. Pattern Recogn 48(5):1773–1796
8. Barron J, Fleet D, Beauchemin S (1994) Performance of optical flow techniques. Int J Comput Vision 12:42–77
9. Mae Y et al (1994) Optical flow based realtime object tracking by active vision system. In: Proceedings of the 2nd Japan-France congress on mechatronics, pp 545–548
10. Neri A, Colonnese S, Russo G et al (1998) Automatic moving object and background separation. Signal Process 66:219–232
11. Stauffer C, Grimson WL (2000) Learning patterns of activity using real-time tracking. IEEE Trans Pattern Anal Mach Intell 22:747–757
12. KaewTrakulPong P, Bowden R (2003) A real time adaptive visual surveillance system for tracking low-resolution colour targets in dynamically changing scenes. Image Vis Comput 21:913–929
13. Ohta N (2001) A statistical approach to background subtraction for surveillance systems. In: Proceedings of the 8th IEEE international conference on computer vision, vol 2, pp 481–486
14. Mittal A, Paragios N (2004) Motion-based background subtraction using adaptive kernel density estimation. In: Proceedings of the 2004 IEEE computer society conference on computer vision and pattern recognition CVPR'04, vol 2, pp 302–309
15. Mikic I, Cosman P, Kogut G, Trivedi M (2000) Moving shadow and object detection in traffic scenes. In: Proceedings of the 15th international conference on pattern recognition, vol 1, pp 321–324

Design and Analysis of Missile Control Effectiveness, Visualized through MATLAB

Kasigari Prasad, P. Keerthana, S. Firoz, G. Akhila, and D. Bandhavi

Abstract Over the past three decades, there have been numerous studies in the field of missile guidance and control. The result has been great progress and a few ways to deal with this problem have emerged. The basic problem is to set the target with great accuracy in an uncertain and noisy environment. One of the first modes of direction that is proposed is to direct the line of sight and navigation evenly, this involves establishing a line of sight between the tracking sensor and the target. This work investigates the problem of design direction and control of a typical air-to-air location capturing a given target using a variety of set-of-direction instructions which are LOS-based and PN-based guidance. The effectiveness of these guidelines is assessed against a given target in terms of the missed miss-distance and the closest time. This paper provides an effective solution to maintain the stability of the missile by calculating the gain margin and phase margin which are effective methods for understanding the concepts of stability.

Keywords Missile guidance · Line-of-sight (LOS) · Proportional navigation (PN) · Air-to-air missile · Stability

1 Introduction

In the age of technological advancement, the most important thing for the national government to consider is to provide security for the people there and to protect them from foreign invasions. Engineers play a major role in the study, design, and development of defense systems.

Some government agencies involved in defense programs are DRDO, ISRO, RAW, etc. Organization for Defense Research and Development (DRDO) is the Research and Development Unit of the Ministry of Defense, Government of India, with a vision for empowerment of India with better defense technology and the goal of gaining confidence in key security technologies and programs [1].

K. Prasad (✉) · P. Keerthana · S. Firoz · G. Akhila · D. Bandhavi
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, Kadapa, India
e-mail: kpd@aitsrajampet.ac.in

The Department of Defense (DoD) is responsible for the security, safety, maintenance, integration, and transport of nuclear weapons. The stated functions of the DoD are necessary to protect workers and property from any potential danger or danger to the nuclear weapon. The DRDO is India's largest defense research organization founded in 1958. The role of DRDO is very sensitive and confidential. DRDO's activities include involvement in the development of defense technologies that cover a wide range of disciplines, such as aeronautics, electronics, combat vehicles, and engineering systems, and metals, arrows, advanced computer simulation, special equipment, navigation systems, information systems, and agriculture [2].

The DRDO operates in several areas such as

- Industry interface
- Technical management.

    (a) Technology transfer
    (b) Acquisition of technology
    (c) Exports of military products.

- To act, in accordance with the Minister of Foreign Affairs, as the main coordinating agency of the Ministry of Defense in all matters relating to foreign government compliance tools related to the acquisition of technology.
- Design and implementation of scientific research programs and design, development, testing, and evaluation, in areas related to national security.
- To direct and manage the agencies, laboratories, institutions, scope, resources, programs, and projects of the Department.
- All matters relating to certification of aircraft design, their equipment, and stores.
- Support the scientific analysis and participation in the acquisition and testing of all weapons systems and related technologies proposed to be acquired by the Department of Defense.
- Arrangements with universities, research institutes, or organizations working abroad for the provision of foreign bursaries and the training of Indian scientists and specialists under the control of the Department.

1.1 Missiles

A missile is a long-range weapon that flies long distances through the air and explodes when we reach our destination. An example of a missile is a rocket. Missiles are of many kinds; the basic classification is guided missile and unguided missile. The unguided missile is any rocket that is aimed manually and fired at a target with the general assumption that it will hit near the target. Targeted missiles work by tracking the target location in space in certain ways (e.g., using radar), chasing it down, and hitting it accurately. Targeted missile systems can be of various types, operating for different operating purposes [3]. Missiles are also known as guided missiles or guided rockets.
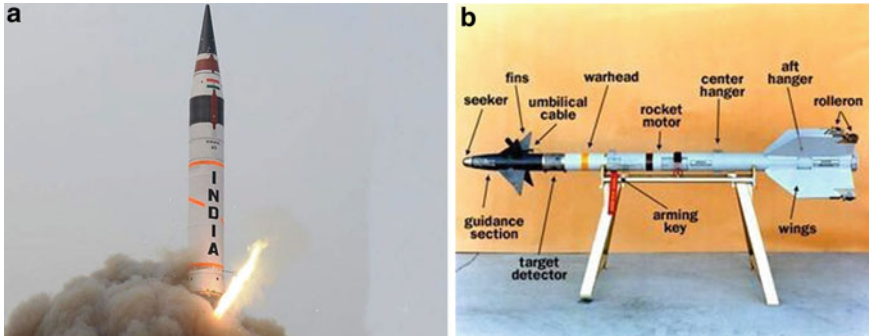
Fig. 1 **a** Representation of missile. *Photo Courtesy*: *The Economic Times*. **b** Components of missile. *Photo Courtesy*: *US Defense System*

Figure 1a depicts a basic representation of the missile, and Fig. 1b depicts the components. The components of the missiles are explained below as

Target detector—The main application of this target detector is to provide information to an arrow aimed at its target or target. It most closely follows any of these technologies such as radar signal, cables, lasers, GPS [4].

Guidance section—The guidance phase is used to process information obtained from the seeker and calculate the correct course of the missile. The steering department (guidance section) is responsible for providing direction to the arrow to achieve its goal or objective.

Warhead—A warhead is an explosive material that is used in military applications to destroy enemy vehicles or buildings [5]. The types of warheads are explosive, chemical, biological and nuclear.

Propulsion system/rocket motor—The propulsion system is to provide the power used to accelerate the missile body and the support needed to achieve the desired target. The propulsion of a missile is available with the help of a rocket engine. It produces thrust by releasing a hot gas called a propellant [6].

Seeker—The seeker is responsible for absorbing the infrared light from the target.

Flight system—The flight system uses data from the pointing or direction system to direct an arrow in an aircraft, allowing it to withstand the accuracy of an arrow or to track a moving target.

Fins—The plane turns on its own, directing the arrows into the air. The wings have been shown solely to give a sense of proportion. The wings have been shown solely to give a sense of proportion.

1.2 Problem Statement

The title of this project is "Investigation of the Design of Missile System and Control." Here, we have estimated the stability conditions of a missile by considering the gain margin and phase margin. This estimated stability is useful to maintain the missile in a stable position to reach the target. These estimated stabilities help reduce the uncertainties in the missile.

1.3 Objectives

The objectives involved in the "Investigation of the Design of Missile System and Control" are

- To study and investigate the working of the missiles
- To maintain the stability of the missile to reach its goal or target
- To estimate the range for obtaining the target kill procedure
- To understand and analyze the launch-to-intercept scenario of the missile which carries out in two phases. They are

  1. Launch phase
  2. Mid-course guidance phase.

2 Literature Review

S. D. Pavan Kumar and Ashline George have proposed a method of using positional servo system in missile system. They have explained the importance of missiles and launch vehicles in the aerospace industry, which is also very important to have a proper model for design and control purposes [7].

Naigang Cui et al. explained a system to control the formation of aircraft arrows. They focus on two aspects such as the flight control arrow and the Back-To-Turn (BTT) intelligence missile outer loop. Where the process of designing an aircraft missile is described in order to produce a single command, as like to be controlling missile [8].

Venugopal Reddy B et al. have proposed architecture to improve the efficiency of integration using efficient resources and providing flexibility and flexibility with minimal space, weight, and power. Authors have developed the concept of inexpensive integrated navigation, navigation and control, and the telemetry system [9].

Gopisetty Srinivas has proposed a laser-guided missile which is a projectile airborne to reach and destroy the target. Here, he has explained that the whole missile body is propelled to hit a predetermined target directed through the laser beam pointing toward the target. In this method, he also proposed and explained that

if the target changes its position, the camera which is present over the missile body will track the location and the updated information is fed to an onboard computer of the missile [10].

Ching-Fang Lin explained about the performance of present-day missile systems which are seriously degraded in reaching the target. Lin also mentioned that usage of classical missile guidance and control techniques is not sufficient enough in defeating these highly maneuvering targets. So Lin proposed and expected new innovations to improve the advancements in missile guidance and control [11].

In the study of numerous papers, several sophisticated control methods have been proposed to improve the basic navigation algorithm for a specific purpose such as shorter time, less power, impact angle, stability, and impact time. Based on this idea, we have considered stability as the main factor in reaching the target. So we have proposed a method to eliminate the uncertainties and errors that occur in the missile after its launch phase. These uncertainties are estimated and eliminated by the proper calculation of the phase margin and gain margin.

The design procedure is illustrated in block diagram of Fig. 2.
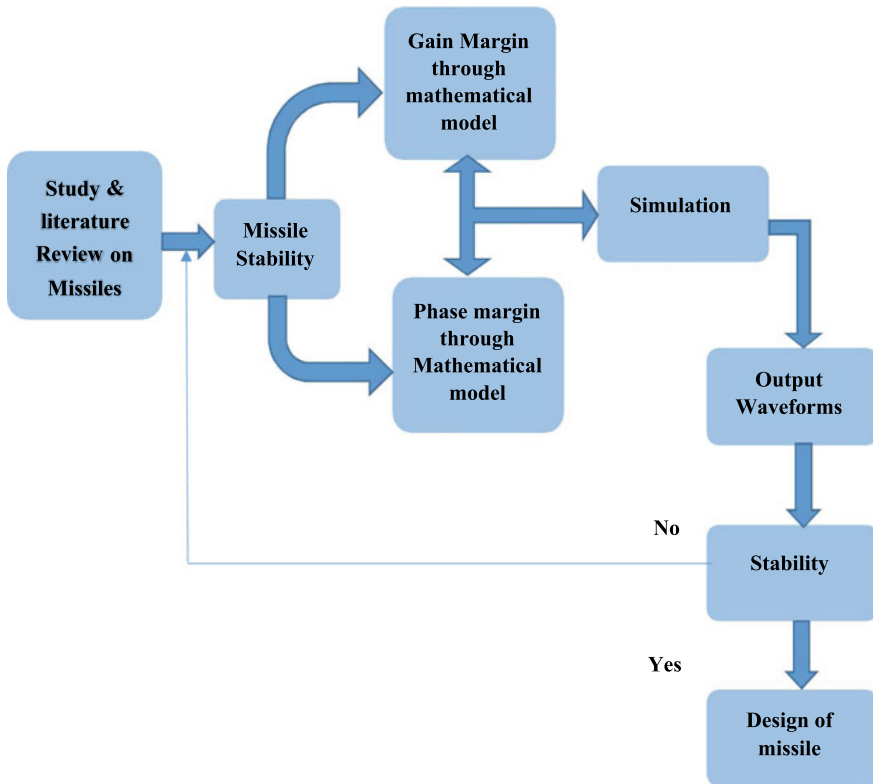
Fig. 2 Block diagram of the procedure to maintain the missile in stable condition

In this paper, the main objective is to maintain the missile in a stable position. The procedure is explained as the investigating or designing a missile involves analyzing the requirements to reach the target, based on the requirements the phase margin and gain margin are calculated by considering the mathematical models and closed-loop transfer functions. The next step involves the simulation of the resultant mathematical equations of the phase margin and gain margin through the MATLAB code; the output signal waveforms are obtained for gain margin and phase margin. Based on this output waveforms of the phase margin and the gain margin and by analyzing them the stability of the missile can be observed, if the obtained output waveforms are matched with theoretically calculated equations of phase margin and gain margin then accordingly the missile will be designed, if the obtained equations are not matched with the theoretically calculated equations then the phase margin and gain margin equations should be calculated again.

3   Equations

For Pitch Axis

Equation supports for calculating the phase margin and gain margin of inner loop

$$\frac{a_z}{a_{zcmd}} = \frac{aU(PK_Ps + PK_I)\left[Z_{\delta e}s^2 + \left(M_{\delta e}Z_q - M_qZ_{\delta e}\right)s - FP\right]}{s.\det1(s) + aU(PK_Ps + PK_I)\left[Z_{\delta e}s^2 + \left(M_{\delta e}Z_q - M_qZ_{\delta e}\right)s - FP\right]}$$

where

a_z = lateral acceleration

Det = determinant equation (calculated for the characteristic equation of transfer function of q (pitch rate) and a$_z$)

a = angle of attack

PK$_p$ and PK$_I$ = gains with respect to the pitch controller U = X-component of velocity

N =Proportional navigation constant s = Laplace operator

FP =perturbed acceleration transfer function in pitch axis

For Yaw Axis

Equation supports for calculating the phase margin and gain margin of outer loop

$$\frac{a_z}{a_{zcmd}} = \frac{aU(PK_Ps + PK_I)\left[Z_{\delta e}s^2 + \left(M_{\delta e}Z_q - M_qZ_{\delta e}\right)s - FP\right]}{\begin{array}{c} s.\det1(s) + (aUYK_PY_{\delta r}s^3 + \left(aUYK_IY_{\delta r} - aUYK_P(N_rY_{\delta r} - N_{\delta r}Y_r)s^2\right) \\ +(aYK_PUFY - aUYK_I(N_rY_{\delta r} - N_{\delta r}Y_r))s + aUYK_IFP] \end{array}}$$
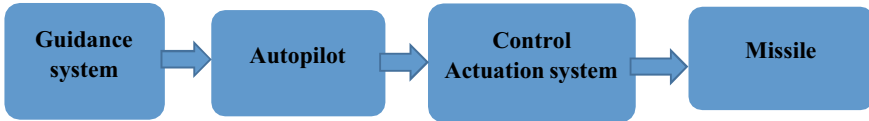
Fig. 3   Block diagram for the operation of the missile

where

YK_P and YK_I = gains with respect to the yaw controller FY = perturbed transfer function in yaw axis

4   Operation of Missiles

The operation of missiles is explained below:

The principle working of missiles is based on Newton's third law, i.e., action and reaction are equal and opposite. Action is caused by producing thrust in the engine and the opposite reaction will be the gases emitting from the back end of the missile.

Figure 3 shows the basic blocks include missile functioning guidance system, autopilot, control actuation system, and missile.

Guidance System: The directional system is a visual or visual device, or a set of devices used to control the movement of a ship, aircraft, arrows, rocket, satellite, or any other moving object. The guidance system improves the accuracy of the arrows by improving its leadership opportunities. These guiding technologies can usually be divided into several sections, with the broadcast sections active, passive, and preset guide.

Autopilot: The automatic flight system controls the flight without the pilot using the controls directly. Such a plan is designed to reduce the workload of human pilots to reduce their fatigue and reduce operating errors during long flights.

Control Actuation System: An actuator is a part of a machine that is responsible for moving and controlling a system or system. Control operating systems are designed to meet extreme accuracy and high performance.

Missile: A rocket-propeled weapon is designed to deliver an explosive warhead with great accuracy at high speed [12].

5   Results and Discussion

In the Fig. 4, the layout shows the time response of the pitch controller by inserting a step. From this we can say that the system is stable and after that, we should get the

phase margin and gain margin both internal and external loop present on the pitch control.

Based on the results obtained as depicted in Table 1, the missile system takes some time to maintain its stability by avoiding all the disturbances occurring in the external environment. That time is known as settling time.

From the Figs. 5 and 6, the phase margin and the gain margin are shown which are plotted by using the commands in the MATLAB from which those are calculated. From this, we can say that the pitch channel is stable.
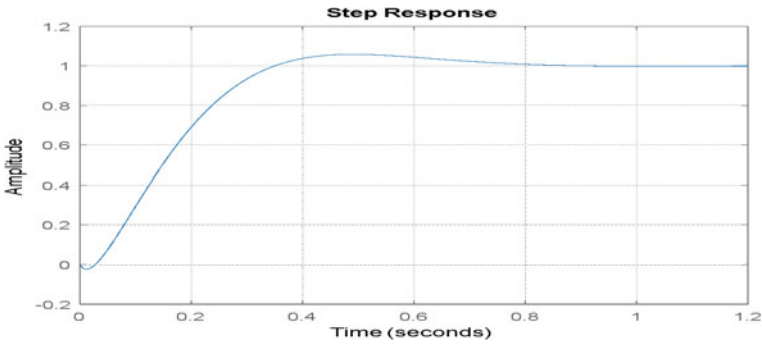
Fig. 4 Step response of pitch channel

Table 1 Analyzed results from the step response of pitch channel

Parameter	Initial value	Final value
Settling time	0	0.4
Stability maintenance time	0.5	1
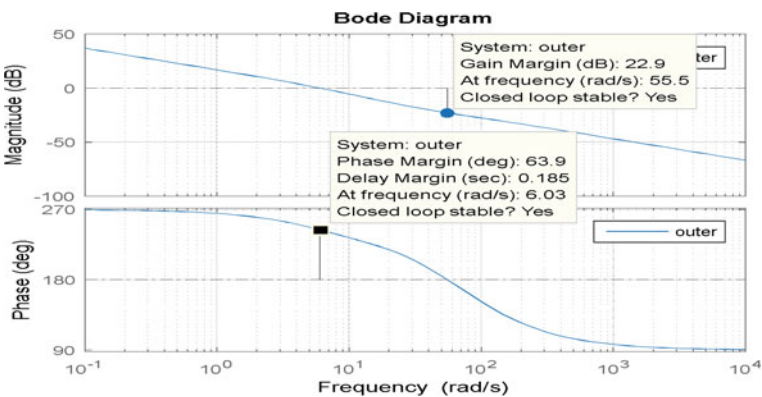
Fig. 5 Phase margin and gain margin of the outer loop
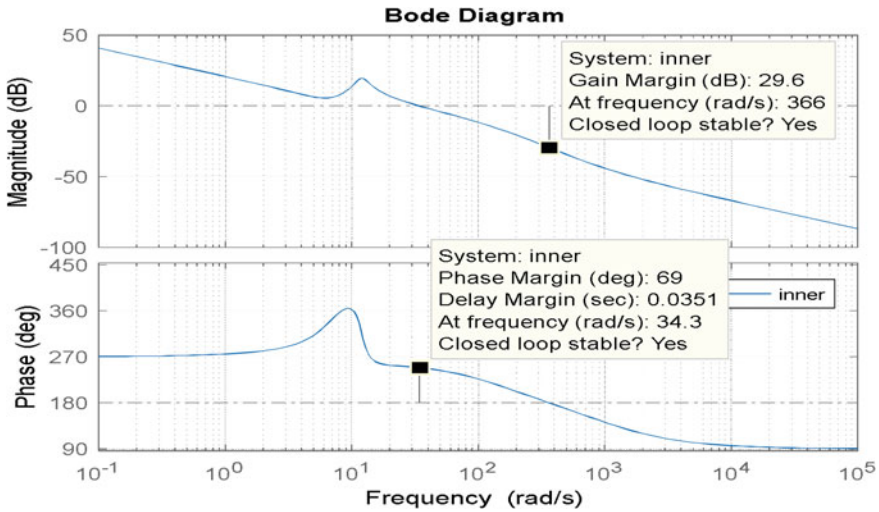
Fig. 6 Phase margin and gain margin of the inner loop

Table 2 Analyzed results from the Bode diagrams of outer loop and inner loop of pitch channel

Parameter	Outer loop	Inner loop
Gain margin	22.9 dB	29.6 dB
Phase margin	63.9 deg	69 deg
Delay margin	0.185 s	0.0351 s

Based on the above results of Table 2, we can conclude that the missile system is maintained in a stable position with the gain margin of 22.9 dB in the outer loop and 29.6 dB in the inner loop. Similarly, a phase margin of 63.9° in the outer loop and 69° in the inner loop is maintained.

6 Conclusion

We have proposed an approach to decrease uncertainty in the interception of the target with great accuracy in the environment by calculating gain margin and phase margins to maintain the stability of the missile. Gain cross-over frequency and phase cross-over frequencies are calculated and considered for operating conditions. The proposed design methodology can be utilized for the design of an autopilot system for all tactical missiles like air-to-air missiles and surface-to-air missiles.

References

1. Speakman EA (1952) Research and development for national. Defense 40(7):772–775
2. Lubas DG (2017) Department of defense system of systems reliability challenges. In: IEEE 2017 annual reliability and maintainability symposium (RAMS), Orlando, FL, USA (2017.1.23–2017.1.26), pp 1–6
3. Vinoth MS, Saradhi PSRP, Aditya PSR (2009) Intelligent guided missile. In: 2009 International conference on IEEE intelligent agent and multi-agent systems (IAMA 2009), Chennai, India (2009.07.22–2009.07.24), pp 1–4
4. Carrera EV, Lara F, Ortiz M, Tinoco A, Leon R (2020) Target detection using radar processors based on machine learning. In: IEEE 2020 IEEE ANDESCON, Quito (2020.10.13–2020.10.16), pp 1–5
5. Xie B, Wang J (2014) Recognition of warhead in ballistic missile defense. In: 2014 7th International congress on image and signal processing, (2014.10.14–2014.10.16) (CISP), Dalian, China, pp 1110–1114
6. Ehsani M, Rahmann KM, Toliyat HA (1996) Propulsion system design of electric vehicles. In: Proceedings of the 1996 IEEE IECON, 22nd international conference on industrial electronics, control, and instrumentation, Taipei, Taiwan, 5–10 Aug, vol 1, pp 7–13
7. Kumar SDP, George A (2018) Missile system modelling for control application. In: 2018 Second international conference on electronics, communication and aerospace technology (ICECA)
8. Cui N, Wei C, Guo J, Zhao B (2009) Research on missile formation control system. In: 2009 International conference on mechatronics and automation
9. Reddy BV, Venkatamani T, Kannan M, Reddy GV (2020) Integrated guidance, navigation, and control system for tactical missile applications. In: 2020 AIAA/IEEE 39th digital avionics systems conference (DASC)
10. Srinivas G et al (2021) Survey on laser guided missile systems and implementation by developing a laser guidance system. Glob J Electron Commun Res 12(1):1–9. ISSN: 2249-314X
11. Lin C-F (1983) New missile guidance and control technology [panel disc.Introduction]. In: 1983 American control conference
12. Gurav B, Economou J, Saddington A, Knowles K (2017) Multi-mode electric actuator dynamic modelling for missile fin control. MDPI

An Unsupervised Sentiment Classification Method Based on Multi-level Sentiment Information Extraction Using CRbSA Algorithm

Shiramshetty Gouthami and Nagaratna P. Hegde

Abstract Social big data is a feedback data or text obtained from the users on social media like Twitter, YouTube, etc. This is large in size and unstructured data format. So there is a need of collecting the information from the massive social big data to the marketers and investors. In order to do this operation, an automatic process is required. There is an enormous change in the social big data every day, so unsupervised techniques are not acquiring the label training data, and it gives the focus on the natural language processing and sentiment classification. Hence, the contrast rule-based sentiment analysis (CRbSA) algorithm is introduced in order to find required information from the massive data by using an unsupervised sentiment classification method. This method works based on "multi-level sentiment" data collecting. This is a starting level of computing the sentiment intensity of users or reviewers. Then after a multi-level union, action based on the sentiment category is introduced. Contrast rule-based sentiment analysis algorithm is developed as a new sentiment algorithm, and it is used to collect or gather the required information automatically. At last, the extraction of required data from the complex social big data can be achieved by the classifier. This is the best method over the baseline in sentiment classification as general counting and SentiStrength algorithms to get the more accuracy. Marketing system uses this type of algorithms to extract the user's feedbacks.

Keywords Social big data · Unsupervised sentiment classification · CRbSA algorithm

S. Gouthami (✉)
Osmania University, Hyderabad, India
e-mail: gouthami.shiramshetty@gmail.com

N. P. Hegde
Vasavi College of Engineering, Ibrahimbagh, Hyderabad, India
e-mail: nagaratnaph@staff.vce.ac.in

1 Introduction

Sentiment classification for documents has recently attracted a lot of investors for its significant in marketing research, public opinion survey, etc. The basic task of sentiment classification is to classify the opinions of public whether it is positive or negative. This is a value of product or services given to the businessmen, and it also gives the idea about product and how the end consumer feels. But there are some problems to know the performance of the sentiment classification. One of them is labeled training data which is needed to perform the supervised data [1]. If it is performed by third party, then it is expensive, and when the data is labeled by the hand, then it takes more time to finish the task. Other one is topical context of a document which cannot be considered into account, so the variation of topical context is large. Generally, the topical context of the document can be taken into account in order to get the higher accuracy. Then there are two steps: first one is the consideration of topical context of the given document and second one is the precision of the algorithm increased for the different sentiment. Because of its model, human labeled data is highly dependent. On the other side, unsupervised data starts its process by collecting the list of sentiment words, e.g., some root words and dictionary words, to find the sentiment words [2]. This paper will represent the new model of unsupervised technique to overcome the three main drawbacks of present methods of unsupervised sentiment classification [3]. This paper is proposing a multi-criteria fusion and multi-level fuzzy computing. By using the multi-level fuzzy, the confusion over the polarity can be solved in the pre-classification stage of sentiment classification [4]. Lack of topical context and confusion over the polarity can be solved by the multi-level criteria fusion in the self-learning stage of sentiment classification.

2 Sentiment Analysis Methods

There are three types of techniques used to separate the opinions of consumers. They are machine learning approach, natural language processing (NLP), and lexicon-based approach and hybrid approach [5, 6]. The two classifications, namely support vector machine (SVM) and Naïve Bayesian are included in the machine learning approach [7, 8]. The information of parts of speech and Word Net can be used by the NLP and lexicon-based approach.

2.1 Machine Learning-Based Approach

This is a fully automatic processing method, and it is having the capability to maintain the large data also, so this approach is more efficient to collect the opinions

of consumers toward the product or service. Supervised, unsupervised, and semi-supervised learning are the three methods of machine learning-based sentiment classification [9].

Supervised Learning The method which has been adopted, detected, and investigated for getting more accurate results from opinions is supervised learning. It is a traditional method or classification for getting the more successful results. Support vector machine (SVM), Naïve Bayes, and K-nearest neighbors (KNN) are the three algorithms used in the supervised classification [10]. In Naïve Bayes, prior probabilities of $P(X|Y)$ and $P(Y)$ are calculated by the generative classifier from the given data, and next it produces posterior probability of $P(Y|X)$ from the prior probabilities bases. In SVM, there are no prior assumptions from the training data, so posterior probability of $P(Y|X)$ is calculated directly by a discriminative classifier. Finally, K-nearest neighbors (KNN) lazy algorithm does not require prior construction. So Naïve Bayes and SVM algorithms are the effective in supervised learning.

Unsupervised Learning The unlabeled training data is collected easily by the text classification, and it is somewhat hard to collect the labeled data. So this difficulty is achieved by the unsupervised learning methods. LDA and pLSA are the conventional models for unsupervised methods which are going to extract the topics from the text documents [11, 12]. To get results efficiently, an unsupervised approach needs more volume of data, so this is one of the drawbacks. Unsupervised models do not give coherence of human analysis with the objective functions. Except this disadvantage, unsupervised models give more information about data without any disturbance.

Semi-supervised Learning (SSL) SSL is a new technique to extract the information by reading the unlabeled data also in real applications [13]. The main idea of SSL is that there is no information of classes in unlabeled data, and it contains only the joint distribution on features classification. The supervised learning learns only the labeled data, whereas both the labeled and unlabeled data can be learned by the SSL. In order to get the best results from supervised learning, it is merged with SSL technique. Then the unlabeled data is improved in the target.

2.2 Lexicon-Based Approach

The text can be analyzed by using opinion lexicon, and this can be named as lexicon-based approach. By using statistical or semantic methods, we can find the words from the specific context. Two methods are present in this approach. First one is dictionary-based approach [14]; in this, the root words can be analyzed, and then it searches the root words and their meanings and opposite meanings in the dictionary. Another one is corpus-based approach [15]; in this, the opinion words are found after starting with list of root words.

3   Multi-level Fuzzy Computing with CRbSA Sentiment Classification

By using multi-level sentiment information extraction, the overall proposed work of unsupervised sentiment classification method by CRbSA algorithm is shown in Fig. 1. The work is classified into three types; first type is lexicon sentiment which works based on unsupervised sentiment classification; second one is to find the sentiment intensity of feedbacks of the multi-level computing method. And the last one is fuzzy classifier estimation and its parameters estimation.

Fig. 1 Framework of multi-level fuzzy computing-based CRbSA sentiment classification algorithm
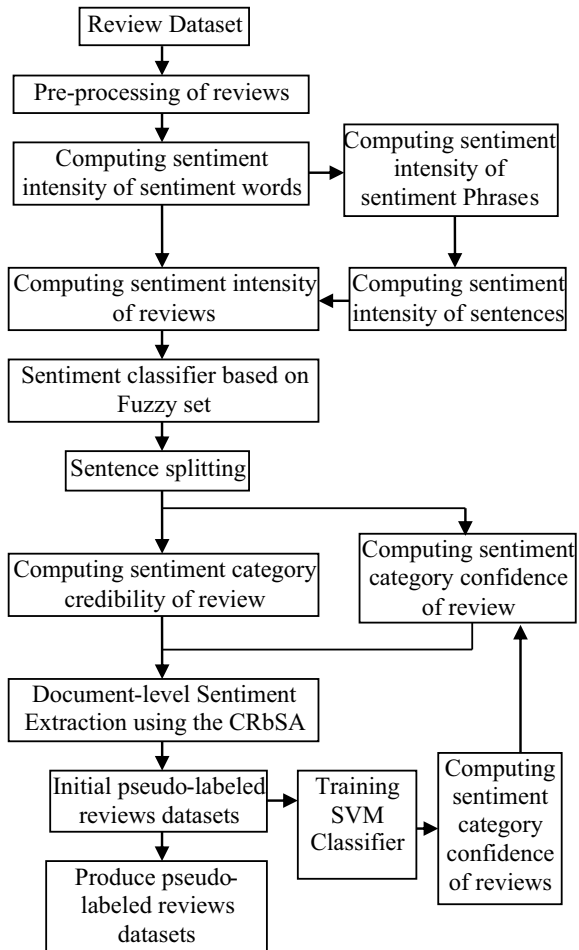
3.1 Sentiment Intensity Calculating Methods on the Word and Phrase Level

By using grammar rules and a fixed-length sliding window, a sentiment phrase structure is designed by taking the reference of sentiment words of candidate. In the sliding window, changing the adverbs gives the calculation about intensity of the sentiment phrase which is done. By using sliding window of length 5 to the limit, the specific method includes sentiment words and discovering of the sentiment phrases and its length with specific rules.

The rules are given as follows:

Adjectives	Adverb + adjective + adverb
Verbs	Adverb + verb + adverb
nouns	Adjective + noun + adverb

The influence of adverbs in sentiment phrases is taken into account while calculating the sentiment phrase intensity, and the representation of Eq. (1) gives the sentiment intensity $si(p_k)$ of sentiment phrase p_k.

$$si(p_k) = f(\text{class(adv)})si(w_k) \tag{1}$$

Here,

$$f(\text{class(adv)}) = \begin{cases} 1, & \text{class(adv)} = \text{na} \\ \alpha, & \text{class(adv)} = \text{da} \end{cases} \tag{2}$$

Here, class(adv) is adverb type, α which indicates weighting factor of degree adverb, na is negative adverb, and degree adverb is da. The degree of adverbs is classified into five types, for adverb types, and each category uses the weighting coefficients as 5β, 4β, 3β, 2β, and 1β as extreme/most, very, more, slightly, and insufficient, respectively. Here, in each weigh β is same and has a value of 0.4.

3.2 Computing Methods of Sentiment Intensity on the Sentence Level

By taking the various sentences and their relationships, the sentence sentiment intensity is calculated in this sentence level. The calculation needs certain steps as follows:

(1) By joining of the sentiment words sentiment intensity and the sentiment phrases sentiment intensity, total sentence sentiment intensity is calculated.

(2) The three types of sentences based on the punctuation are declarative, inter-
rogative, and exclamatory. Different sentences follow the different methods
for computing the intensity. Sentence types such as declarative, interrogative,
exclamatory sentences, and their respective processing methods are direct calcu-
lation of sentence sentiment intensity, reversing of sentence sentiment intensity,
weighting the sentence sentiment intensity with coefficient of 2.

(3) The relation difference between sentences is based on the adverbs connected
in the sentences. The two main relations of the sentences are adversative and
summary relations. To find the two types of relations, different methods are
adopted. The total intensity of the review is calculated by addition of all
sentences sentiment intensities in the review.

3.3 Sentiment Classifier Based on Fuzzy Set

The fuzzy sets are used to define the sentiment categories in classifying the reviews
when the language is fuzzy and especially the sentiment intensity fuzziness is present.
Equation (3) represents the positive sentiment category of the review which is $R = \{r_i\}$ as fuzzy set P,

$$P = \left\{ \left(r_i, \ \mu_p(r_i) | r_i \in R \right) \right\} \tag{3}$$

Here, for the review r_i the member function is $\mu_p(r_i)$, and r_i is belongs to positive
sentiment category P. Equation (4) represents the semi-trapezoid function as the
member function of review r_i,

$$\mu_p(r_i) = \begin{cases} 0, \ si(r_i) < \alpha \\ \frac{si(r_i)-\alpha}{\beta-\alpha}, \ \alpha \leq si(r_i) \leq \beta \\ 1, \ si(r_i) > \beta \end{cases} \tag{4}$$

Here, the sentiment intensity of the review r_i is $si(r_i)$, and adjustable parameters
are α and β which compute the member function boundary. The negative sentiment
category is also defined as the definition of positive sentiment category. Equation (5)
defines the negative sentiment category review set $R = \{r_i\}$ as fuzzy set N,

$$N = \{(r_i, \ \mu_N(r_i) | r_i \in R)\} \tag{5}$$

Here, for the review r_i the member function is $\mu_N(r_i)$, and r_i is belongs to negative
sentiment category N. Equation (6) represents the semi-trapezoid function as the
member function for review r_i,

$$\mu_N(r_i) = \begin{cases} 1, \ si(r_i) < \alpha \\ \frac{si(r_i)-\alpha}{\beta-\alpha}, \ \alpha \leq si(r_i) \leq \beta \\ 0, \ si(r_i) > \beta \end{cases} \tag{6}$$

Here, the sentiment intensity of the review r_i is $si(r_i)$, and adjustable parameters are α and β which compute the member function boundary. Depending on the principle of maximum membership, the result is unified fuzzy set after the addition of positive and negative member functions of the fuzzy set. Equation (7) represents classification functions of the fuzzy sets,

$$
\begin{aligned}
f_k(si(r_i)) &= \max\{\mu_p(r_i), \ \mu_N(r_i)\} \\
&= \begin{cases} r_i \in N, \ si(r_i) \leq \frac{\alpha+\beta}{2} \\ r_i \in P, \ si(r_i) \leq \frac{\alpha+\beta}{2} \end{cases}
\end{aligned}
\tag{7}
$$

Here, the sentiment intensity for the review r_i is $si(r_i)$, the member function for the fuzzy set P is $\mu_p(r_i)$, and the member function for the fuzzy set N is $\mu_N(r_i)$.

3.4 Splitting Sentences

The machine is unable to read the dataset of review which is in the unstructured format. By using the regular expression method, the data is splitting into sentences that mean the string content into string sentences at starting level. This process gives the improvement in finding the sentiment words in the review data, at starting the separation of the data from sentences is done by using punctuations. Those punctuations are question mark (?), exclamatory mark (!), and full stop (.). After that the sentence can be compounded by using conjunctions and, so on, which, what, etc.

3.5 Computing Sentiment Category Credibility of Review CM LC ($\mathbf{R_m}$)

In the self-learning level, the present methods use the initial pseudo-labeled datasets. By these datasets, the sentiment classifier is trained. Then the sentiment reliability of reviews can be calculated by using the sentiment classifier. Based on this reliability of sentiment category, the initial pseudo-labeled review data is added to the selected part of pseudo-labeled reviews. Other than these self- learning methods, the initial pseudo-labeled review data is going to add to the domain category credibility $C(r_m)$, high sentiment category representation with review $R(r_m)$. The multi-criteria fusion strategy is as follows:

1. For any review, r_m belongs to $RD = \{r_m | m = 1, 2, \ldots, N\}$, by depending on the multi-level fuzzy computing model $C_{MLFCM}(r_m)$, the sentiment category of reliability of the review is calculated. This computing processing is shown in Eq. (8).

$$C_{\text{MLFCM}}(r_m) = \begin{cases} \frac{si(r_m)}{si(r_N)}, & si(r_m) \geq 0; \\ \frac{si(r_m)}{si(r_1)}, & si(r_m) < 0; \end{cases} \tag{8}$$

2. To generate initial pseudo-labeled review dataset, the selection process of reviews is high absolute values of $C_{\text{MLFCM}}(r_m)$ and $R(r_m)$, which must be greater than $R_{av}(RD^T)$ (in this method, 40% of the size of the initial pseudo- labeled review dataset is selected).

3.6 Document-Level Sentiment Extraction Using the CRbSA Algorithm

Based on the difference, the sentiment of the sentence is calculated. In the given example, the word "but" in front clause decreases the positive or negative sentiment and in the back clause it increases the negative or positive sentiment. Overall sentiment of a sentence is back clause sentiment. To calculate the total sentiment of the document, there is a need to know the sentiment sentences. To extract the sentiment information of an entire document, a new algorithm is proposed, that is, the contrast rule-based sentiment analysis (CRbSA) algorithm. This algorithm is reprocessed the data and extracted the sentence level. Later the document holding positive and negative sentences percentage can be computed by this algorithm. Based on the average percentage of positive sentences, the sentiment information of a document is calculated. The rules of sentiment extraction are given in Table 1.

1. Initial pseudo-labeled review dataset is trained with the self-learning sentiment classifier (in this process, SVM classifier is selected) for review r_m, the credibility $C_{\text{SVM}}(r_m)$ is calculated by trained classifier, and it is not in the pseudo-labeled review dataset. This process is shown in the equation format as in Eq. (9).

$$C_{\text{SVM}}(r_m) = \begin{cases} \frac{w^T r_m + b}{\max(w^T r_m + b)}, & w^T r_m + b \geq 0 \\ \frac{w^T r_m + b}{\min(w^T r_N + b)}, & w^T r_m + b < 0 \end{cases} \tag{9}$$

Here, from the review r_m to the classifier hyperplane, the distance is $w^T r_m + b$ to the classifier hyperplane.

Table 1 Rules to extract sentiment based on the average percentage of sentence-level positive

Percentage of positive	Sentiment information
Average positive > 55%	Positive
Average positive < 45%	Negative
Average positive ≥ 45% Average positive ≥ 55%	Neutral

2. The calculation of overall reliability $C(r_m)$ of r_m by the $C_{\text{MLFCM}}(r_m)$ and $C_{\text{SVM}}(r_m)$.

$$C(r_m) = \frac{C_{\text{MGFC}}(r_m) + C_{\text{SVM}}(r_m)}{2}$$

3. High values of pseudo-labeled reviews $C(r_m)$ and high values of $C(r_m)$ are to be selected to produce the pseudo-labeled dataset S, and it can be done by combining the $C(r_m)$ and $R(r_m)$.

$$S = \{r_m\}$$

4. The original pseudo-labeled review dataset is added to the reviews in S when the S is not empty.
5. Reviews are added to the original pseudo-labeled review dataset, until no new reviews are there to add. The steps (1)–(4) are repeated until no new reviews to be found. A pseudo-labeled review trains dataset, and a trained sentiment classifier is achieved.

4 Results

The first dataset is noisy social media data, and the second dataset is movie reviews. 25,000 mostly popular movie reviews are there in dataset with positive and negative reviews. The first dataset uses Sentiment 401 that is having the positive, negative, and neutral labels as tweets. Sentiment of a word depends on the context in the sentiment classification, and then it is going to be a challenge for movie review dataset. When relating the movie's area, many words can be used in the horror type which is assumed to be negative in the objective type. While classifying a sentiment, the topical context of a word is used. The main aim is to do the sentiment classification of the review; in this process, the three metrics are evaluated; those are accuracy (AC), recall (R), and $F1$ and precision (P). Those are expressed as below:

$$P = \frac{1}{2}\left(\frac{R_1}{R_1 + W_2} + \frac{R_2}{R_2 + W_1}\right)$$

$$R = \frac{1}{2}\left(\frac{R_1}{R_1 + W_1} + \frac{R_2}{R_2 + W_2}\right)$$

$$F1 = \frac{2PR}{P + R}$$

$$AC = \frac{R_1 + R_2}{W_1 + W_2 + R_1 + R_2}$$

Here, R_1 is true positive, R_2 is true negative, W_1 is false positive, and W_2 is false negative. The true positive is identifying positive sentiment over the positive sentiment of datasets. The false positive is identifying positive sentiment over the negative sentiment of datasets. The false negative is identifying negative sentiment over the positive sentiment of datasets. The true negative is identifying negative sentiment over the negative sentiment of datasets.

The methods SentiStrength, general word counting, CRbSA, multi-level fuzzy computing, and proposed multi-level fuzzy computing with CRbSA sentiment classification algorithm are evaluated. In this paper, the above measures have been used and then calculation of the respective accuracy measure by testing two datasets has been considered. Table 2 shows performance comparison of different sentiment classification algorithms in terms of the R_1, R_2, W_1, and W_2 with their respective accuracy.

Figure 2 shows the accuracy comparison of such various sentiment classification algorithms. From these result analyses, it was said that the proposed sentiment classification framework achieves higher accuracy than the other classification algorithms.

Table 2 Performance comparison of different sentiment classification algorithms

Sentiment classification algorithms	R_1,	R_2	W_1	W_2	AC (%)
SentiStrength	57	29	71	43	43
General word counting	62	28	72	38	45
CRbSA	62	65	35	38	63.5
Multi-level fuzzy computing	77	74	23	26	80.2
Multi-level fuzzy computing + CRbSA	85	87	15	13	86.4

Fig. 2 Accuracy comparison
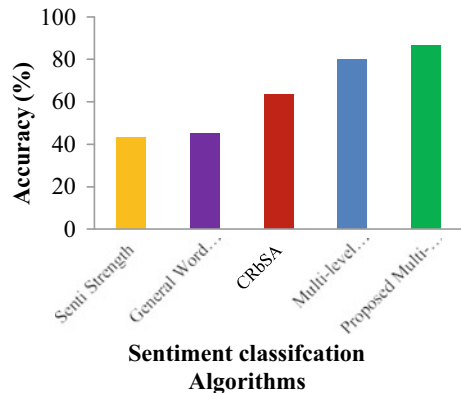
5 Conclusion

This paper presented an unsupervised sentiment classification method by using the CRbSA algorithm which is based on the multi-level sentiment analysis. Even the lack of context data and the uncertainty in the polarity of the sentiment then the fuzziness of the data is solved by the multi-level fuzzy computing method in that especially pre-classification stage of the sentiment analysis. Multi-level sentiment analysis uses the CRbSA algorithm to collect the required data from the unstructured data of reviews. For the extraction of the information by using rules and different negations to examine the sentences especially more complex data, the CRbSA algorithm reflects the details about the complication of the human language. So CRbSA algorithm gives the better results and accurate performance in the negative sentiments than the other sentiment analysis. Two datasets experimental outputs showed the better performance results in this proposed system. This improves the output results of online reviews classification methods while comparing other current methods.

References

1. Zhou L, Zhang Y, Tian Y, Fan W, Zhang T (2021) Deep cross-modal face naming for people news retrieval. IEEE Trans Knowl Data Eng 33
2. Yu S, Niu J, Wang L (2020) SentiDiff: combining textual information and sentiment diffusion patterns for twitter sentiment analysis. IEEE Trans Knowl Data Eng 32
3. Wang W, Cheng X, Meng D, Lin Z, Xu X, Jin X, Wang Y (2016) An unsupervised cross-lingual topic model framework for sentiment classification. IEEE/ACM Trans Audio Speech Lang Process 24
4. Cowie R, Parthasarathy S, Busso C (2016) Using agreement on directions of change to build rank-based emotion classifiers. IEEE/ACM Trans Audio Speech Lang Process 24
5. Zimmer F, Sabetzadeh M, Arora C, Briand L (2015) Automated checking of conformance to requirements templates using natural language processing. IEEE Trans Softw Eng 41
6. Trancoso I, Wong DF, Chao LS, Zeng X (2015) Graph-based lexicon regularization for PCFG with latent annotations. IEEE/ACM Trans Audio Speech Lang Process 23
7. Laface P, Cumani S (2014) Large-scale training of pairwise support vector machines for speaker recognition. IEEE/ACM Trans Audio Speech Lang Process 22
8. Huang D, Yang F, Gao X, Shang C (2014) Novel Bayesian framework for dynamic soft sensor based on support vector machine with finite impulse response. IEEE Trans Control Syst Tech 22
9. Ostendorf M, Wu W (2013) Graph-based query strategies for active learning. IEEE Trans Audio Speech Lang Process 21
10. Tang C-K, Li D, Chen Q (2013) KNN matting. IEEE Trans Pattern Anal Mach Intell 35
11. Kadobayashi Y, Pang S, Kasabov NK, Ban T (2012) LDA merging and splitting with applications to multiagent cooperative learning and system alteration. IEEE Trans Syst Man Cybern Part B (Cybernetics) 42
12. Huang X, Yu X, An A, Liu Y (2012) Mining online reviews for predicting sales performance: a case study in the movie domain. IEEE Trans Knowl Data Eng 24
13. Gómez-Chova L, Bovolo F, Muñoz-Marí J, Camp-Valls G, Bruzzone L (2010) Semi-supervised one-class support vector machines for classification of remote sensing data. IEEE Trans Geosci Remote Sens 48

14. Mishra P, Basu K (2010) Test data compression using efficient bitmask and dictionary selection methods. IEEE Trans VLSI Syst 18
15. McLean D, Li Y, Crockett K, O'Shea JD, Bandar ZA (2006) Sentence similarity based on semantic nets and corpus statistics. IEEE Trans Knowl Data Eng 18

Timestamp Prediction Using Enhanced Adaptive Correlation Clustering-Based Recommender System for the Long Tail Items

Soanpet Sree Lakshmi, T. AdiLakshmi, and Bakshi Abhinith

Abstract This is the study of recommender systems. Timestamps for the items in the long tail are predicted based on the way the items are clustered. The long tail Items have lesser ratings and rarely appear in the recommendations in contrast to the popular items. This work divides the items into two parts according to their popularities, head part and the tail part namely. Long tail items are adaptively correlation clustered. The head part items are grouped according to their popularity. In this paper, the authors propose the prediction of the timestamp of the long tail items using the Enhanced adaptive correlation-based clustering.

Keywords Long tail problem · Adaptive clustering · Correlation clustering · Timestamp prediction · Recommender systems

1 Introduction

The items like movies, music, news are recommended by recommender systems (RS) to their users. These recommendations improve user experience. RS helps organizations to enhance their businesses by recommending items to a targeted cluster of customers [1]. Different types of issues were addressed by RS and the Authors of [2–4] listed many of such issues. The RS is based on item ratings. The RS gives more preference to popular items which have a high number of ratings. The items having a lower number of ratings than popular items are called niche items. On the contrary, in most datasets, the number of items that are popular are less when compared to the long tail items [5]. Long tail items have great potential to improve sales in business [6]. Personalization of recommendations can be done to improve the quality of recommendations. Different approaches are proposed to improve recommendations

S. S. Lakshmi (✉)
Department of Information Technology, Vasavi College of Engineering, Hyderabad, India
e-mail: s.sreelakshmi@staff.vce.ac.in

T. AdiLakshmi · B. Abhinith
Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, India
e-mail: t_adilakshmi@staff.vce.ac.in

of the items in the long tail [7–9]. This work enhances the quality of recommendation by predicting the timestamps of the items rated. The adaptive clustering method proposed in [9] and [10] forms the basis of this work. The method divides the items based on their number of ratings into head and tail sets. The head items are recommended based on the conventional each item Method. The items in the tail part are then clustered based on their popularity. Different types of recommendations based on time-aware approaches were discussed in [11–15]. The recommendations were thus enhanced by predicting the timestamps that would increase the usability of the recommendation.

We propose timestamp prediction based on an enhanced recommender approach focusing on clustering on the correlation of the long tail items.

2 Related Work

2.1 Recommendation Systems

The authors in [4, 16–19] and [20] discussed various types of recommender systems classified to be, content-based, collaborative-based, knowledge-based and hybrid models. Amatriain et al. [4] discusses the various data mining approaches used for recommendation systems. Clustering is one of those suitable approaches.

2.2 Adaptive Clustering

The method of clustering used in [9] outperformed the conventional approach by applying different recommendation approaches in the head and tail parts. Therefore, a similar approach to the one described in [9] is enhanced for recommending items in the long tail by employing a different clustering technique.

2.3 Correlation-Based Clustering

The clustering method proposed in [21] is applied the authors proposed correlated connected clustering to improve clustering accuracy. In [21], density-based clusters were generated and the next point in the neighborhood is included in the cluster only if it has desired correlation with the core point.

Authors in [22] show the identification of cluster representatives using correlation-based approaches.

2.4 Enhanced Adaptive Correlation-Based Clustering

The authors in [10] emphasized a method where the computation overhead reduced in head part recommendation is reduced and the tail part performance is improved by adaptive correlation-based clustering method.

2.5 Time Aware Recommendations

The authors in [11–15] discuss various methods of applying time awareness in recommendations. Most of these papers emphasize on the usable time slice in recommendations. The focus of our approach is predicting time classes to improve the quality of recommendations.

3 Proposed Methodology

The proposed method of predicting time stamps in recommendations system is based on enhanced adaptive- correlation-clustering. This paper implements the adaptive approach. Long tail items are correlation clustered according to their content-based features. The timestamp of the rating was considered and then the Time-class of the movie in test data is predicted. This enhances the quality of the recommendation.

The recommendations are enhanced by predicting the timestamps that would increase the usability of the recommendation. This work proposes the prediction of timestamps. The recommendation quality is enhanced by predicting the time classes for the long tail items, i.e., the niche movies with a lower count of ratings. We propose Timestamp prediction using an enhanced approach focusing on correlation clustering of the long tail items.

We applied the technique to the dataset and the accuracy of the Time-class prediction is studied.

3.1 Dataset

The Movielens [23, 24] data contains 20,000,263 ratings and 465,564 tag applications across 27,278 movies. This dataset provides movie related tags information and the ratings given by a user to different movies. 138,493 users rated the movies on a scale of 1–5. These ratings are provided. The tag genome data structure encodes the content-based features namely tag relevance scores for movies. These values describe how strongly movies exhibit relevance to each tag. The method proposed is tested on part of this dataset.

A subset of 1000 users' data was considered. The ratings provided by these users are used in the experiment. Minimum criteria of 10 ratings for a movie is specified. 25,154-ratings, timestamp prediction accuracy for 2895 movies is discussed in results.

The Elbow method is used to split the data into two parts.

The rating count of movies is sorted, and the plot of sorted ratings is presented in Fig. 1.

The data in ratings.csv appears as in Fig. 2.

Every rating in the dataset has an attribute "timestamp". This attribute has been converted to a proper date and time for further classification. The date and time are classified into 8 different clusters. 2-bit representation is used to identify the classes using the method below:

(a) The first bit was set to 1 if it was a weekend and 2 for weekday
(b) The second bit was set to either of {1, 2, 3, 4} depending on the rating timestamp as shown in Table 1.

Fig. 1 Rating distribution for the dataset considered (1000 Users)
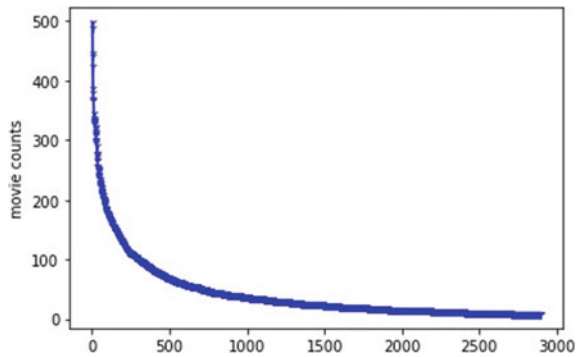
Fig. 2 Ratings.csv

1	userId	movieId	rating	timestamp
2	1	2	3.5	1.11E+09
3	1	29	3.5	1.11E+09
4	1	32	3.5	1.11E+09
5	1	47	3.5	1.11E+09
6	1	50	3.5	1.11E+09
7	1	112	3.5	1.09E+09
8	1	151	4	1.09E+09
9	1	223	4	1.11E+09
10	1	253	4	1.11E+09
11	1	260	4	1.11E+09

Table 1 Time classes classification

Second bit	Time period
1	3.00–9.00
2	9.01–15.00
3	15.01–21.00
4	21.01–3.00

3.2 The Derived Variables

The following are the variables defined for each user and movie.

(a) Using the count of the movies with greater than average ratings, user wise favorite clusters are identified
(b) Using the maximum values of ratings given by the users, the user favorite genre is calculated as User favorite genre.
(c) Using the average values of the ratings given by the users, the user genre average for each genre is calculated.
(d) Using the number of ratings each movie gets, movie rating count is computed
(e) Using the values of ratings each movie gets from all the users, movie average rating is calculated
(f) Each movie is categorized into one of the many genres. It is described by the movie genre variable.

3.3 Flow Chart

Figure 3 shows the flow chart of the proposed method of timestamp prediction.

3.4 Algorithm

Step 1: Every rating in the dataset has an attribute "timestamp". This attribute has been converted to a proper date and time for further classification.

$$(1.11E + 09 \text{ converted to Sat Mar 5 } 10 : 50 : 00 \text{ 2005})$$

Step 2: The date and time are classified into 8 different clusters. 2-bit representation is used to identify the classes using the method below:

(a) The first bit was set to 1 if it was a weekend and 2 for weekday
(b) The second bit was set to either of {1, 2, 3, 4} as specified in Table 2
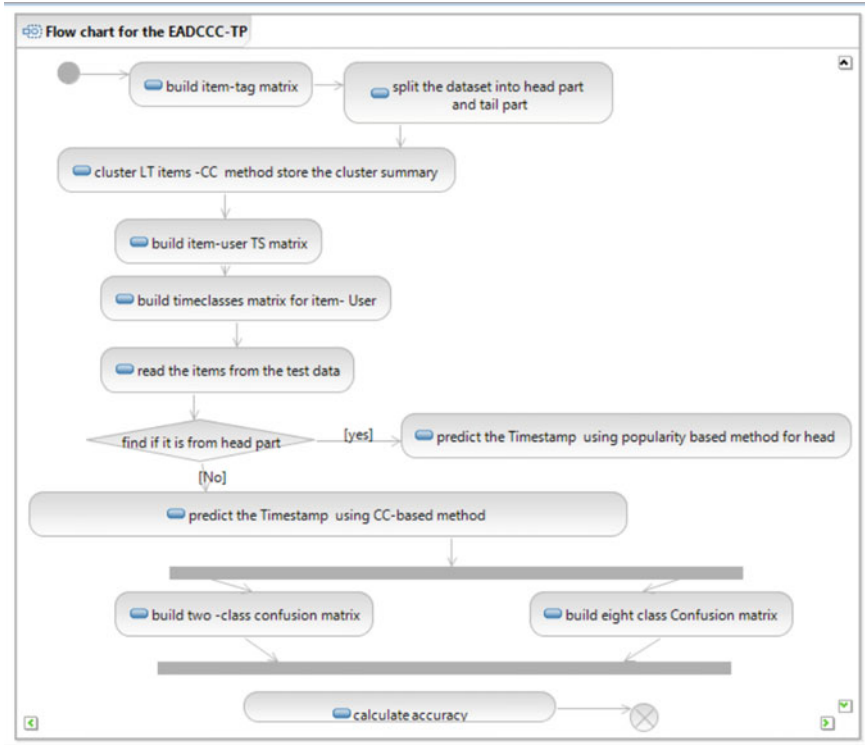
Fig. 3 Flow chart of the proposed methodology

Table 2 Two-class confusion matrix for time-class prediction

	Actual weekend (P)	Actual weekday (N)
Predicted—weekend	195 (TP)	202
Predicted—weekday	857	2565 (TN)

Step 3: user_movie_timestamp_matrix that consisted of timestamps for every rating is created.

Step 4: Following are the sub-steps to find the predicted timestamp for each movie:

I.   Movie's genre is identified and stored into movie_genre.
II. The movie's clusterId was taken into consideration and stored into clus_id.
III. Genres of all the movies in the cluster of clus_id were extracted.
IV. If there exist more movies of the same genre in the cluster then

(a) All such movies are extracted.
(b) All the timestamp_cluster of those movies are extracted
(c) Maximally occurring timestamp_cluster is given as a prediction.

Else:

(a) Most occurring genre in the cluster is taken
(b) All movies with that genre are extracted
(c) All the timestamp_cluster of those movies are extracted
(d) Maximally occurring timestamp_cluster is given as a prediction.

Step 6: Generate a confusion matrix from the predictions made.

4 Experimental Results

The confusion matrix for weekend and weekday prediction for sample test data is presented in Table 2.

4.1 Two-Class Confusion Matrix for Time-Class Prediction

4.2 Eight- Class Confusion Matrix for Time-Class Prediction

Confusion matrix for the weekend and weekday prediction with predicted time-class for sample test data is presented in Fig. 4.

Sample Accuracy calculation for Class 11.

$$= (TP + TN)/(TP + TN + FP + FN)$$

P\\A	11	12	13	14	21	22	23	24	Accuracy%
11	27	9	2	9	11	12	12	17	92.08
12	1	5	5	2	3	1	2	3	95.21
13	0	0	3	2	3	1	1	1	95.6
14	26	7	10	87	37	21	21	56	85.54
21	47	33	34	62	249	78	81	148	75.37
22	18	11	15	41	50	108	28	66	84.78
23	8	9	12	26	42	21	69	47	86.11
24	130	97	82	232	311	218	220	827	57.35

Fig. 4 Eight class confusion matrix for time-class prediction

$$= (27 + 3488) \big/ (27 + 3488 + 72 + 230)$$
$$= 3515 \big/ 3817$$
$$= 0.9208$$
$$= 92.08\% \text{ Average accuracy for all 8 time classes} = 84.005\%$$

5 Conclusion

In this paper, we predicted the timestamps using the enhanced adaptive method for long tail items and the accuracy is studied. The timestamp prediction assumes that the user provided the timestamp at the time the movie is watched. However, the accuracy of the prediction can further be improved if user-related data like age, gender, occupation were provided.

References

1. Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: a survey of the state-of-the art and possible extensions. IEEE Trans Knowl Data Eng 17(6):734–749
2. Lakshmi SS, Adi Lakshmi T (2017) The survey of recommender systems. Int J Eng Trends Technol (IJETT), Special Issue—April 2017
3. Park DH, Kim HK, Choi IY, Kim JK (2012) A literature review and classification of recommender systems research. Expert Syst Appl 39:10059–10072
4. Amatriain X, Jaimes A, Oliver N, Pujol J (2011) Data mining methods for recommender systems. In: Ricci F, Rokach L, Shapira B, Kantor PB (eds) Recommender systems handbook. Springer US, pp 39–71
5. Cremonesi P, Garzotto F, Pagano R, Quadrana M (2014) Recommending without short head. In: Proceedings of the 23rd international conference on world wide web (www) '14 companion, April 7–11
6. Anderson C (2006) The long tail. Hyperion Press
7. Yin H, Cui B, Li J, Yao J, Chen C (2012) Challenging the long tail recommendation. In: Proceedings of the VLDB endowment, vol 5, no 9
8. Park YJ, Tuzhilin A (2008) The long tail of the recommender systems and how to leverage it. In: Proceedings of the ACM conference on recommender systems, pp 11–18
9. Park YJ (2013) The adaptive clustering method for the long tail problem of recommender systems. IEEE Trans Knowl Data Eng 25(8)
10. Lakshmi SS, Adi Lakshmi T, Abhinith B (2021) An adaptive correlation clustering-based recommender system for the long-tail items. In: Smart innovation, systems and technologies, vol 224. Springer, pp 505–514. http://www.springer.com/series/8767
11. Shi Y (2014) An improved collaborative filtering recommendation method based on timestamp. In: 16th International conference on advanced communication technology. IEEE
12. Sun L, Michael EI, Wang S, Li Y (2016) A time-sensitive collaborative filtering model in recommendation systems. In: 2016 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData), Chengdu, pp 340–344

13. Yuan Q, Cong G, Ma Z et al (2013) Time-aware point-of-interest recommendation. In: Proceedings of the international ACM SIGIR conference on research and development in information retrieval, ACM, Dublin, Ireland, July 2013
14. Wei S, Ye N, Zhang Q (2012) Time-aware collaborative filtering for recommender systems. In: Proceedings of the Chinese conference on pattern recognition, Sept 2012. Springer, Beijing, China
15. Huang Z, Stakhiyevich P (2021) A time-aware hybrid approach for intelligent recommendation systems for individual and group users. Complexity 2021(8826833):19. https://doi.org/10.1155/2021/8826833
16. Pazzani M (1999) A framework for collaborative, content-based and demographic filtering. Artif Intell Rev 13:393–408
17. Adomavicius G, Tuzhilin A (2011) Context-aware recommender systems. In: Ricci F, Rokach L, Shapira B, Kantor PB (eds) Recommender systems handbook. Springer US, pp 217–253
18. Schafer JB, Frankowski D, Herlocker J, Sen S (2007) Collaborative filtering recommender systems. In: Brusilovsky P, Kobsa A, Nejdl W (eds) The Adaptive web. Springer, Berlin Heidelberg, pp 291–324
19. Pazzani M, Billsus D (2007) Content-based recommendation systems. In: Brusilovsky P, Kobsa A, Nejdl W (eds) The adaptive web. Springer, Berlin Heidelberg, pp 325–341
20. Burke R (2000) Knowledge-based recommender systems. In: Encyclopedia of library, and information systems, vol 69, pp 175–186
21. Bohm C, Kailing K, Kroger P, Zimek A (2004) Computing clusters of correlation connected objects. In: Proceedings of the ACM international conference on management of data (SIGMOD), Paris, France, pp 455–466
22. Nagy D, Aszalós L, Mihálydeák T (2019) Finding the representative in a cluster using correlation clustering. Pollack Periodica Int J Eng Inf Sci 14(1):15–24 https://doi.org/10.1556/606.2019.14.1.2
23. http://files.grouplens.org/datasets/movielens/ml-20m-README.html
24. http://files.grouplens.org/papers/tag_genome.pdf

In Cloud Computing Detection of DDoS Attack Using AI-Based Ensembled Techniques

Alka Shrivastava and Pratiksha Gautam

Abstract The growing interest in cloud computing has resulted in an increase in the number of cyber-attacks counter to it. One such attack is a Distributed Denial of Service (DDoS) attack, which targets the cloud's B/W, resources, and services in order to render them inconvenient to both their customers and cloud supplier. Because of the large volume of traffic that must be processed, machine learning classification algorithms, and data mining have been suggested to distinguish common packets from anomalous packets in order to improve efficiency. When it comes to cloud DDoS attack defense, feature selection has also been analyzed as an initiation phase that has the potential to increase classification accuracy, while simultaneously decreasing computational complexity by analyzed most significant features from the actual dataset, which is done in the time of supervised learning. In this paper, we supposed an ensemble-based multi-filter feature selection techniques with together the o/p of four different filter techniques in order to execute the best possible selection. An extensive experimental evaluation of our suggested technique was carried out used to the intrusion detection benchmark dataset, the NSL-KDD, and a decision tree classifier, among other tools. If we compare the results obtained with those obtained using other classification techniques, we can see that our suggested method successfully decreasing the No. of features from 41 to 13, and it has classification accuracy with high detection rate.

Keywords Feature selection · Cloud computing · Filter methods · DDoS · Machine learning · AI

A. Shrivastava (✉) · P. Gautam
Department of CSE, Amity School of Engineering and Technology, Amity University Gwalior, Gwalior, MP, India
e-mail: alkacse2009@gmail.com

P. Gautam
e-mail: pgautam@gwa.amity.edu

1 Introduction

Cloud computing has exploded in popularity as a result of its ability to offload computation and meet storage needs on them and, depending on the availableness of the user's computer. In addition to cloud computing on-demand self-service, provides the following notable features [1], rapid elasticity, broad N/W access, measured services and resource pooling as shown in Fig. 1. Depending on the services provided, we can be divided cloud computing into three categories—Software as a Service, Platform as a Service, and Infrastructure as a service, private cloud computing, public cloud computing, and multi-cloud or hybrid cloud computing are all types of cloud computing deployments. As a result of their ability to provide a high-quality service in recent years, Amazon, Google, IBM, and Microsoft are the greatly prominent cloud service providers (CSPs).

In the cloud computing environment, security concerns are major concerns and a difficult task to accomplish. The DDoS attack is particular most serious security matter in a cloud climate [2], and it is particularly dangerous. It is defined as any malicious or event behavior that prevents or reduces a cloud's ability to perform its services and expected functions in order to maintain its capacity. DDoS attacks resulted in downtime, economic loss, and other short- and long-term consequences for the victim CSP's operations. It is more potent than a DDoS attack in this the attacker making a defense force before launching an attack in the class of bots or zombies. All of those bots have been programmed to attack the victim and disable the victim's ability to function in some way. They take advantage of the CSP's property offer and attempt to flood the area. DDoS attacks could be divided into 2 categories: semantic attacks and brute-force attacks.
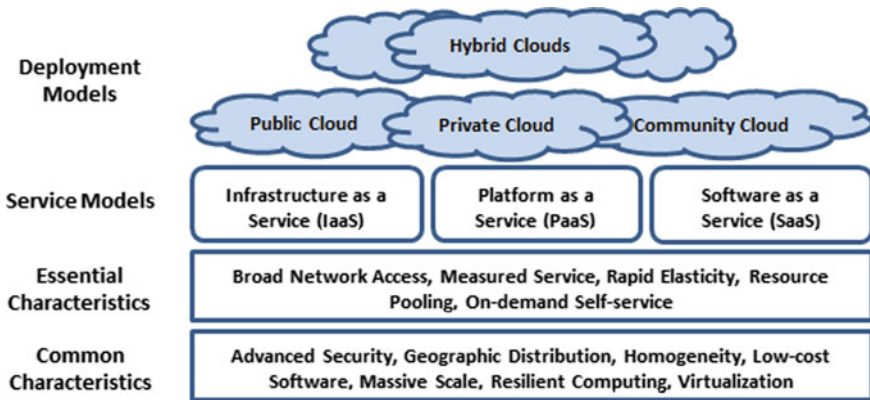

Fig. 1 Architecture of cloud computing

1.1 High Rate DDoS Attack

The high rate or flooding attack, also called as a brute force attack, is a type of cyber-attack. An enormous amount of malicious demand is sent by the attackers in an attempt to disturb the n/w B/W of a particular cloud server. In order to cause a disruption in connectivity, it is necessary to disable the router's processing power and N/W Bandwidth capacity. The N/W-level flooding attack is referred to as a high-rate attack. High-rate attacks include flooding the transmission control protocol (TCP) [3], flooding the user datagram protocol (UDP), and flooding the Internet control message protocol (ICMP). By removing server resources such as disc space, memory, and CPU from the system the unavailability of cloud services to legitimate users is achieved. Attacks on the application level include the hypertext transfer protocol (HTTP) flood attack [4], the simple mail transfer protocol (SMTP) flood, and the domain name system (DNS) flood attack, among others.

The attackers initiate such attacks by make use of the vulnerabilities of a large number of computers conducive to make attack armies, which are collectively referred to as botnets. The attacker could make the control, which is then passed on to the cloud server, which in turn forwards the control to the numerous cooperating hosts in the network. The cooperative hosts forward the influx of requests to further cloud servers that have been designated as targets. The botnet computer system can employ an IP spoofing method concerning launch Distributed DoS attacks in order to conceal the true source of the attack. As a result, determining the genuine position of the attacker is a difficult but necessary assignment.

1.2 Low-Rate DDoS Attack

The semantic attack, also known as a low rate attack or a vulnerability attack, is a typed attack this takes favor of a protocol weakness. The attackers only transmit a small quantity of malicious traffic to the destination at a time. When related to high rate attacks, discovery the lowest rate attack is a very challenging and important task to accomplish. The low-rate distributed DoS attack is more complex and challenging to discover, than the high-rate Distributed DoS (DDoS) attack, owing to the low-rate traffic and silent behavior of the attack. This is due to the fact that the attacker transmits malicious appeal at a very low rate, which is obscured by the traffic volume based defense mechanisms. Instead of terminating cloud services, the attack has an impact on the Quality-of-Service (QoS) that the legitimate user has specified. Low-rate attacks incorporate the shrew attack, the prune of Quality attack, the Low-Rate distributed DoS attack, and the Economic Dos (EDoS) attack [5].

Fig. 2 Life cycle of DDoS attack
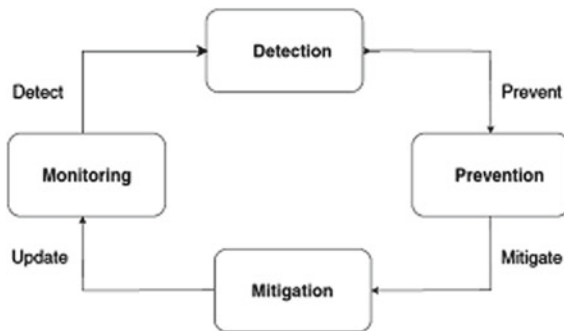
1.3 Life Cycle of DDoS Attack

Figure 2 [6] depicts the four phases of a Distributed DoS (DDoS), which are monitoring, detection, prevention, and mitigation. When performing host monitoring or N\W, it is necessary to record remarkable information about the host or N\W. The detection phase requires examining the N\W traffic that has been captured in order to identify the malicious try. The prevention phase is acclimated protect the resources and cloud service from being misused by developers who wish to develop some applications in multiple locations. In the end, the mitigation phase is based on the calculated severity of the attack and involves taking precise action concerning manage its consequences and ramifications. The outcomes of the mitigation phase are sending on to the prevention phase, where they are used to update the preventive measures in place. This paper focuses solely on the detection phase, which is the most comprehensive of the four phases examined.

1.4 Detection of DDoS Attack

The Distributed DoS attack is an extremely challenging safety matter this causes traffic at the time resource proportion in a cloud climate. As a result, detecting DDoS attacks is a critical task in order to provide competent resource proportion to end users [7]. Detection techniques for Distributed DoS (DDoS) attacks could be greatly categorized into two types, signature-based detection technique [7] and anomaly-based detection technique [8]. The signature-based detection approach captures N/W traffic, which is then related to precise attack patterns, for example, bytes or packet sequences, to determine whether or not an attack has occurred. When related to anomaly-based detection methods, that category of detection scheme is lots simpler to comprehend and develop, and it produces significantly more significant results. The signature-based detection scheme, on the other hand can only detect known attacks where the pattern has already been defined. When an attack is detected, the

anomaly-based detection approach is helped to determine the source of the attack through the use of behavioral patterns.

According to [9] the detection mechanism can be divided into a number of different categories. Identification schemes that are based on virtual machines, pattern matching and fingerprints, replace point detection-based, filtering-based, feature segment based, data mining-based and entropy-based are all possible. The authors of this paper conduct a brief survey on detection methods that make use of artificial intelligence algorithms in this paper. Consequently, many AI approaches are using to detected DDoS attacks in a cloud climate, including (support vector machines (SVMs) [10], (random forests (RFs) [4], naïve Bayes (NBs) [8], decision trees (DTs), artificial neural networks (ANNs) [11], k-nearest neighbors (KNN) classification [12], convolutional neural networks (CNNs), particle swarm optimization (PSO), and k-nearest neighbors with the help of RF ensemble machine learning method and the Information Theoretic Entropy (ITE), authors in [4] have developed a detection system for HTTP Distributed DoS attacks. The entropy (H) of the N/W header features of incoming N/W traffic is calculated using a time-based sliding window approach that is applied over time. The authors presented a method to detect and unknown Distributed DoS (DDoS) and mitigate known attacks in real-time climates by employing an artificial neural N/W (ANN) algorithm considering detect Distributed DoS attacks based on peculiar distinguish features this isolated Distributed DoS attack traffic from legitimate traffic. The adaptive hybrid neuro-fuzzy systems-based detection technique suggested by Kumar et al. for detecting Distributed DoS (DDoS) attacks in a cloud climate [13] is based on neuro-fuzzy systems and adaptive hybrid neuro-fuzzy systems. The proposed NFBoost method obtains the final classification decision by combining an ensemble of classifier O/P with a Neyman Pearson cost reduction methods concerning get the ending classification decision.

S.M. Vardhan et al. suggested a new GOA algorithm that incorporates a GOIDS [14] and a new grasshopper optimization algorithm (GOA). Making an IDS concerning meet the needs of the monitored condition and allow categorize between an attack and normal traffic is the basis for the proposed approach, which is carried out in the following manner. Aside from that, GOIDS (machine learning algorithm) is evoking the applicable features from the actual intrusion detection system (IDS) dataset that could be used to identify the most common low-speed DDoS attacks. Once this is done, the features that were selected are used as I/Ps to the classifiers. The DT, MLP, NB, and SVM are using for identify the attack that has been launched into the network.

Using a very fast decision tree (VFDT) research arrangement in the cloud-assisted Wireless Body Area N/w [15] Suggested a new detection concept for distributed victim-based DDoS (Distributed DoS) attacks in cloud-assisted Wireless Body Area N/W (WBAN). The suggested arrangement improved the accuracy of a Distributed Denial of Service (DDoS) attack while simultaneously lowering the false positive and false negative ratios. Using Taylor-elephant herd optimization (FT-EHO) and an effective fuzzy classifier inspired by the deep belief network (DBN), [16] suggested a latest detection concept for discover the Distributed DoS attack by combining a fuzzy

classifier with a Taylor series and elephant herd escalation algorithm combining along a fuzzy classifier for rules research. [17] Using artificial immune systems to identify the compelling features of an attack, that paper supposed a latest Distributed DoS detecting system that makes avail oneself of artificial immune systems. This supposed detecting technique is able of discover threats and responding in accordance along the behavior of the biological resistance procedure, which is a significant advantage over existing methods. A multilayer perceptrons merge sequential feature selection method is proposed by [18] concerning choose the foremost attribute in the time of the training phase. Afterwards, when the attack detector detects significant detection errors, the feedback structure is implemented to reconstruct the attack detector.

Supposed a latest attack detector based on a probabilistic neural N/W based on PSO (PSO-PNN) [19]. Initialization involves converting the user's behavior into a meaningful and understandable format. In the following step, a multi-layer neural network was used to classify and identify the malicious behaviors, Punitha and colleagues.

Punitha and Indumathi [20] supposed a latest centralized cloud information answerability system; the integrity along imperialist competitive key generation algorithm (CCIAI-ICKGA) is utilized by attackers to hack into the system. Additionally, the supposed technique has the capability of detecting an attack as well as monitoring the real time usage of the users' detail. Cipher text-policy feature-based encryption (CP-ABE) along key generation construct utilizes the ICKGA, and a trapdoor generator is utilized to produce the private and public private keys for each user in the system. The trapdoor generator confirms that user data is accurate and complete both at the cloud server and at the level of the individual user. Ultimately, a dynamically weighted ensemble neural N/W (DWENN) classifier is employed in order to find out the distributed DoS attack along greater sensitivity and accuracy.

With the help of a mathematical model based on queuing theory [21] describe the flow table-space of a switch. The flow-table dividing technique is utilizing to safe the S/W SDN-based cloud from distributed DoS attacks that overload the flow table. The supposed methods aim to rise the effectiveness of the cloud system's defense in case of distributed DoS attacks while requiring the least involvement of the SDN controller. Xu et al. [22] represented a distributed DoS detection techniques establish on K-Means CC and Fast K-Nearest Neighbors that was effective in detecting DDoS attacks (K-FKNN).

2 Methodology

Using the o/p of the 1/3 divided of ranked features from the filter techniques represented above, we have developed our proposed EMFFS method. EMFFS is an initialization phase that takes place previous to learning, during which sole filter approaches are utilized for the basic selection mechanism of candidates. To rank the feature set of the original dataset, the methods IG, gain-ratio, chi-square, and relief F are used. After creating a mutually exclusive subset of the ranked features, a 1/3 break of the
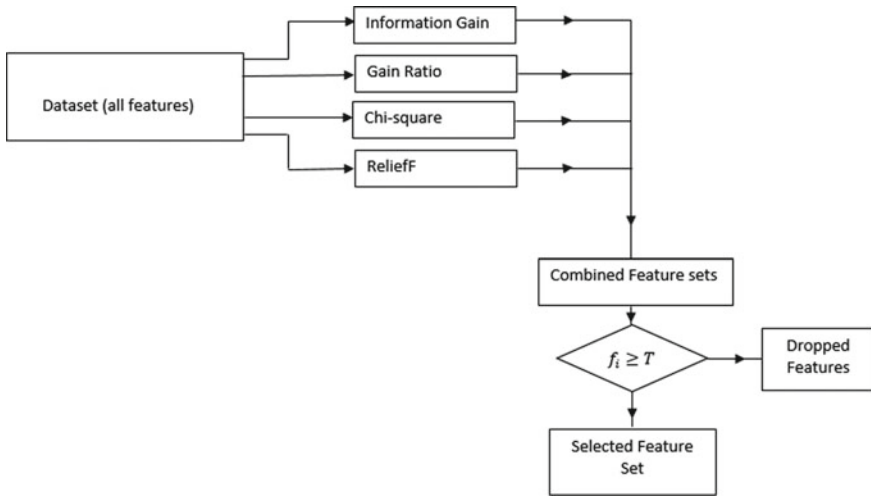
Fig. 3 Flowchart and block diagram of ensemble-based multi-filter feature selection approach

ranked features is selected (i.e., 14 features). Those characteristics are regarded as the paramount characteristics of each filter method in terms of performance. It is decisive by margining the o/p of each filter approaches and used an easy majority vote to decisive the last selected attribute that the EMFFS produces as a result of the combined output of each filter method. Among the four filter methods, a threshold is decisive conducive to recognize the frequently happen attribute, and it is set at three (i.e., $T = 3$). After margining all of the selected attribute sets, a counter is using to decisive which attribute are common across all of the attribute sets in relation to the threshold value. Only those features are selected for further consideration this accommodated the threshold standard and are using as the last attribute set for classification. The supposed EMFFS approach is depicted in Fig. 3.

The EMFFS approach is built by the algorithms describe below:

Algorithm 1.1 (Filter feature ranking approaches)

Step 1: let X_i be the features set in the NSL-KDD dataset,

where $X_i = \{X_1, X_2, X_3,... X_{41}\}$ and C_I describe the class (i.e. normal or anomaly),

where $C_i = \{C_1, C_2\}$

Step 2: For each filter approaches rank and sort the features X_i correspondent

to its significance in decisive the o/p class C_i

Step 3: Select 1/3 break of each filter selection approaches o/p X'.

Algorithm 1.2 (Combine output features)

Step 1: Merge selected o/p features X' of each filter techniques.

Step 2: Find out the feature count threshold T.

Step 3: Calculate the feature occurrence rate in company of the filter techniques.

Algorithm 1.3 (Ensemble choose)

Step 1: Select intercepts of common features from Algorithm 1.2

Step 2: If the feature count is smaller than the threshold, leave the feature or else choose the feature.

Step 3: Repeat step 2for all the features in the 1/3 break subset.

2.1 Classification Algorithm and Dataset

A famous data mining classifier for forecast is the decision tree classification algorithm, which is famous because it is simple to understand the interaction between variables. A greedy algorithm [23] is used to repetitive build a decision tree, with the divide-and-conquer strategy being used to accomplish this. The tree is composing of the internal nodes, root node, leaves, and branches, each of which describe a rule that is required to categories data in the manner of its attributes and is represented by the tree. To division each i/p into each internal nodes in accordance along the attribute of the data record, decision tree using supervised dataset along root node being the first characteristic and test condition to break individual i/p into each internal nodes [24]. In cases where the root node has the greatest information gain, the preceding node along the later greatest information gain is chosen as the test for the later node. It is necessary to repeat that process up to all of the characteristic have been related, or up to all samples belong to the same class and there is no remaining characteristic that can be used to further partition the samples [25]. It is possible to connect two nodes with a branch, as well as a node and a leaf, using a branch. Each node is composed of branches, each of which is labeled with the possible values of the characteristic in the parent node. The decision value of classification is labeled on the leaves of the tree.

Observe a case that was chosen at random from a set S of cases that belonged to the category C_i. In order to determine the probability this arbitrary sample be in to class C_i, the following formula [25] can be used in Eq. 1

$$P_i = \frac{\text{freq}(C_i, S)}{|S|} \tag{1}$$

where $|S|$ is the No. of samples in the set S. So, the information it forward could be described by-log2P_I bits. At this time, propose the probability distribution is stated as $\{P_{1, 2}, \ldots, P_n\}$, accordingly, the information lugged by the distribution, this is entropy of P, could be conveyed in Eq. 2

$$\text{Info}(P) = \sum_{i=1}^{n} -P_i \log_2 P_i \tag{2}$$

Apportionment a set of K samples, based on the value of a non-categorical attribute X, into sets K_1, K_2,…, K_m, the information be in need to decisive the class of an element of K is the weighted average of the information required to analyze the class of an element KI. The weighted average of Info (Ki) could be decisive by Eq. 3.

$$\text{Info}(X,\ K) = \sum_{i=1}^{m} \frac{|K_i|}{K} \times \text{Info}(K_i) \tag{3}$$

The information gain, in(X,), could thus be estimated as follows in Eq. 4.

$$\text{Gain}(X,\ K) = \text{Info}(K) - \text{Info}(X,\ K) \tag{4}$$

Above equation represent the subtraction between the information necessary to spot an element of K and an information required to analyze an element of K after the value of attribute X has been described. Therefore, that is the information gain due to attribute X.

3　Results

For constructing decision tree, different algorithms are used; C5.0 and its earlier version C4.5 has been represented in [26], though, for our method, we will have used J48, a version of C4.5 as our classifier. It's a decision tree classification algorithm based on Iterative Dichotomiser 3. It's very supportive in check out the data continuously and categorically [27, 28] True Negative Rate (TNR) and accuracy are calculated for proposed design with singular value decomposition and without singular value decomposition. We used MATLAB 2018a version to execute the results [29, 30] and the Figs. 4 and 5 show the results.

4　Conclusion

Particular most significant challenges currently overlook by N/W intrusion systems in cloud computing is the operating of massive internet traffic during a DDoS (Distributed Denial of Service) attack. Preprocessing datasets before attack classification in cloud computing has been accomplished through the use of feature selection methods. It has been demonstrated in that work that an ensemble-based multi-filter feature selection techniques can combine the o/p of a 1/3 split of ranked important features such as relief F, chi-squared, gain ratio, and information gain, with the o/p
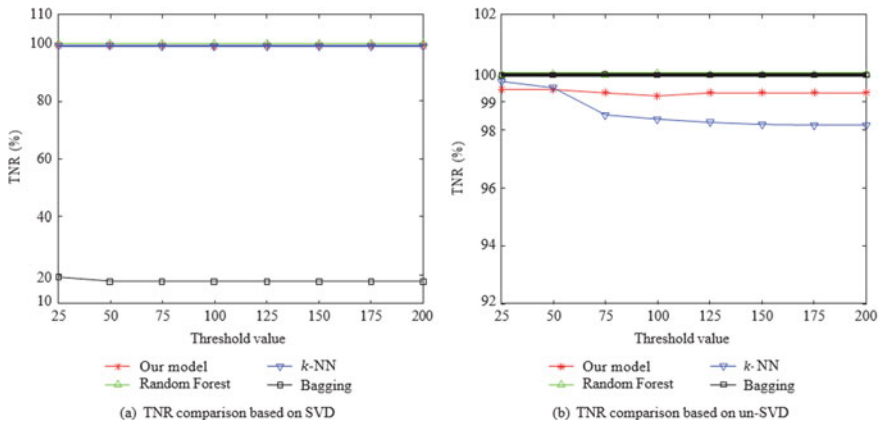
Fig. 4 TNR for comparing proposed model with other algorithms
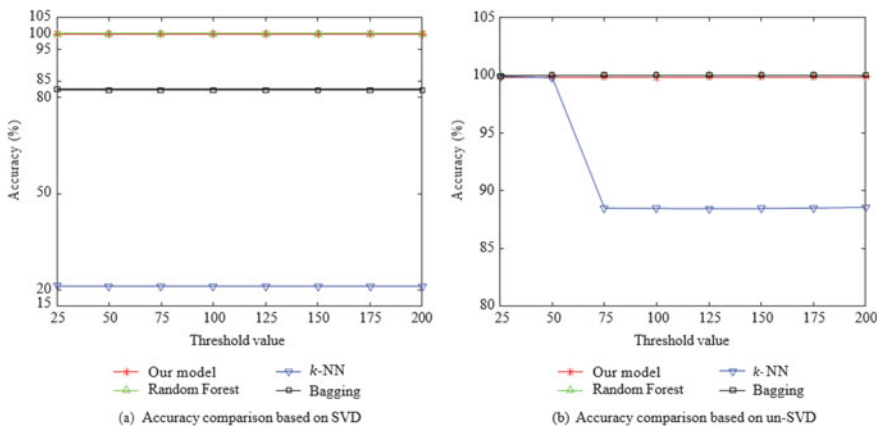
Fig. 5 Accuracy for comparing proposed model with other algorithms

of an ensemble-based multi-filter feature selection method. It is decisive by linking the o/p of each filter method and applying a predetermined threshold to determining the final feature used a simple majority vote that the resulting output of the EMFFS will be. The EMFFS method with 13 features outperforms other filter methods used the J48 classifier and other supposed feature selection techniques when tested on the NSL-KDD dataset, according to the results of the performance evaluation.

References

1. Masdari M, Jalali M (2016) A survey and taxonomy of DoS attacks in cloud computing. Secur Commun Netw 9(16):3724–3751
2. Lee K, Kim J, Kwon KH, Han Y, Kim S (2008) DDoS attack detection method using cluster analysis. Expert Syst Appl 34(3):1659–1665
3. Osanaiye OA, Dlodlo M (2015) TCP/IP header classification for detecting spoofed DDoS attack in cloud environment. In: IEEE EUROCON 2015—international conference on computer as a tool (EUROCON). IEEE, pp 1–6
4. Idhammad M, Afdel K, Belouch M (2018) Detection system of HTTP DDoS attacks in a cloud environment based on information theoretic entropy and random forest. Secur Commun Netw 2018
5. Ghanbari M, Kinsner W (2022) Detecting DDoS attacks using polyscale Analysis and deep learning. In: Research anthology on smart grid and microgrid development. IGI Global, pp 1078–1096
6. Somani G, Gaur MS, Sanghi D, Conti M, Buyya R (2017) DDoS attacks in cloud computing: issues, taxonomy, and future directions. Comput Commun 107:30–48
7. Osanaiye O, Choo KKR, Dlodlo M (2016) Distributed denial of service (DDoS) resilience in cloud: review and conceptual cloud DDoS mitigation framework. J Netw Comput Appl 67:147–165
8. Rawashdeh A, Alkasassbeh M, Al-Hawawreh M (2018) An anomaly-based approach for DDoS attack detection in cloud environment. Int J Comput Appl Technol 57(4):312–324
9. Agrawal N, Tapaswi S (2019) Defense mechanisms against DDoS attacks in a cloud computing environment: state-of-the-art and research challenges. IEEE Commun Surv Tutorials 21(4):3769–3795
10. Polat H, Polat O, Cetin A (2020) Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models. Sustainability 12(3):1035
11. Saied A, Overill RE, Radzik T (2016) Detection of known and unknown DDoS attacks using artificial neural networks. Neurocomputing 172:385–393
12. Amjad A, Alyas T, Farooq U, Tariq MA (2019) Detection and mitigation of DDoS attack in cloud computing using machine learning algorithm. EAI Endorsed Trans Scalable Inf Syst 6(26)
13. Kumar PAR, Selvakumar S (2013) Detection of distributed denial of service attacks using an ensemble of adaptive and hybrid neuro-fuzzy systems. Comput Commun 36(3):303–319
14. Dwivedi S, Vardhan M, Tripathi S (2020) Defense against distributed DoS attack detection by using intelligent evolutionary algorithm. Int J Comput Appl 1–11
15. Latif R, Abbas H, Latif S (2016) Distributed denial of service (DDoS) attack detection using data mining approach in cloud-assisted wireless body area networks. Int J Ad Hoc Ubiquit Comput 23(1–2):24–35
16. Wang M, Lu Y, Qin J (2020) A dynamic MLP-based DDoS attack detection method using feature selection and feedback. Comput Secur 88:101645
17. Velliangiri S, Pandey HM (2020) Fuzzy-Taylor-elephant herd optimization inspired deep belief network for DDoS attack detection and comparison with state-of-the-arts algorithms. Futur Gener Comput Syst 110:80–90
18. Prathyusha DJ, Kannayaram G (2021) A cognitive mechanism for mitigating DDoS attacks using the artificial immune system in a cloud environment. Evol Intel 14(2):607–618
19. Rabbani M, Wang YL, Khoshkangini R, Jelodar H, Zhao R, Hu P (2020) A hybrid machine learning approach for malicious behaviour detection and recognition in cloud computing. J Netw Comput Appl 151:102507
20. Punitha A, Indumathi G (2021) A novel centralized cloud information accountability integrity with ensemble neural network based attack detection approach for cloud data. J Ambient Intell Humaniz Comput 12(5):4889–4900

21. Bhushan K, Gupta BB (2019) Distributed denial of service (DDoS) attack mitigation in software defined network (SDN)-based cloud computing environment. J Ambient Intell Humaniz Comput 10(5):1985–1997
22. Xu Y, Sun H, Xiang F, Sun Z (2019) Efficient DDoS detection based on K-FKNN in software defined networks. IEEE Access 7:160536–160545
23. Gehrke J, Ganti V, Ramakrishnan R, Loh WY (1999) BOAT—optimistic decision tree construction. In: Proceedings of the 1999 ACM SIGMOD international conference on management of data, pp 169–180
24. Sánchez-Marono N, Alonso-Betanzos A, Tombilla-Sanromán M (2007) Filter methods for feature selection—a comparative study. In: International conference on intelligent data engineering and automated learning. Springer, Berlin, Heidelberg, pp 178–187
25. Xiang C, Yong PC, Meng LS (2008) Design of multiple-level hybrid classifier for intrusion detection system using Bayesian clustering and decision trees. Pattern Recogn Lett 29(7):918–924
26. Bujlow T, Riaz T, Pedersen JM (2012) A method for classification of network traffic based on C5. 0 machine learning algorithm. In: 2012 International conference on computing, networking and communications (ICNC). IEEE, pp 237–241
27. Lin SW, Ying KC, Lee CY, Lee ZJ (2012) An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection. Appl Soft Comput 12(10):3285–3290
28. Shrivastava R, Singh M, Teja KSSR (2021) A real-time implementation for the speech steganography using short-time Fourier transformior secured mobile communication. J Phys: Conf Ser 2089(1):012066. IOP Publishing
29. Tiwari R, Sharma M, Mehta KK, Awasthy M (2020) Dynamic load distribution to improve speedup of multi-core system using MPI with virtualization. Int J Adv Sci Technol 29(12s):931–940
30. Tiwari R, Sharma M, Mehta KK (2020) IoT based parallel framework for measurement of heat distribution in metallic sheets. Solid State Technol 63(6):7294–7302

Securing Data in Internet of Things (IoT) Using Elliptic Curve Cryptography

Nagaratna P. Hegde and P. Deepthi

Abstract The data everywhere was being transferred for each second using domain of Internet of Things (IoT). To secure data using the Internet of Things (IoT) is a tedious task. When the data is being transferred using Internet of Things (IoT), more security can be provided using elliptic curve cryptography. Asymmetric cryptography is used by most of the applications for providing secure communication between two parties (Weber in Comput Law Secur Rev 26:23–30, 2010). The purpose of this type of cryptography is the requirement of huge amount of computation and storage. This is where the use of elliptic curve cryptography comes into picture, as it needs less storage and can be used in small computational devices. ECC needs smaller key sizes and provides stronger encryption compared to various asymmetric cryptographic algorithms like RSA. The usage of power required is low but the performance of the devices using ECC is high for different types of devices like IoT, sensors, etc. This paper shows how ECC must be implemented strongly for providing communication securely to encode the data on an elliptic curve, in IoT devices. The encryption of the data must be done securely while mapping the data on to the elliptic curve. This paper shows how the data is encrypted and mapped on to the curve securely. The work shows how the data can be encrypted with ECC and how it can be visible to only authorized users. In the field of cryptography, ECC is a method for asymmetric cryptography which is dependent on the algebraic structure of an elliptic curve on the finite field.

Keywords Asymmetric key cryptography · Internet of Things (IoT) · Elliptic curve cryptography (ECC)

N. P. Hegde
Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, Telangana, India

P. Deepthi (✉)
Computer Science and Engineering, Bhoj Reddy Engineering College for Women, Hyderabad, Telangana, India
e-mail: deepthiputnala@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
A. Kumar et al. (eds.), *Advances in Cognitive Science and Communications*, Cognitive Science and Technology, https://doi.org/10.1007/978-981-19-8086-2_95

1 Introduction

The Internet of Things (IoT) facilitates data exchange in the connected devices, vehicles, software, etc., in a network. The IoT is used for providing the infrastructure for exchanging secure and reliable data. The IoT consists mainly of the integration of sensors, RFID tags and different technologies for communication. The IoT describes how to integrate various physical objects and devices with the Internet so that they can work together and exchange data with one other for achieving some common goals. The IoT is primarily made up of small, interconnected materials to facilitate the context of collaborative computing. IoT limits include connectivity and computing power [2].

Life has become easy with the help of IoT devices, so little attention should be paid for providing security to the devices. Currently, developers are focusing on ways to improve the capabilities for the devices along with considering the security of the devices. Data sent over IoT networks is not secure enough. The data transferred over IoT must be protected to provide privacy for the users. If there is no security of data over IoT, data exchanged may be breached and also the personal information can be hacked. The main intention of IoT is to provide authentication, confidentiality and integrity. Authentication must be provided when the communication is taken place over the network, without authentication hackers can easily communicate with any of the devices.

The data is transferred between two devices when they communicate with each other. The data that is exchanged may be very sensitive or personal, moving around devices over the IoT network, which must be encrypted. To protect the data from intruders, one needs to encrypt the data [3]. The data exchanged can be encrypted using cryptography for processing the text in unreadable form. The goal of encryption is to provide integrity, confidentiality, authentication and non-repudiation. In the work presented, ECC is used for encryption. One of the public key cryptographies is ECC which depends on algebraic structure of elliptic curve on a finite field.

This paper shows how the data is transferred using an elliptic curve in a secured way [4]. The method used here uses some specific ASCII values for encryption or decryption of the data. The process of encryption and decryption is done smoothly. The method used reduces the size of cipher text for more number of words. ECC provides better security with smaller key sizes and is suitable for devices with less power and storage.

Cryptography is used for converting plain text into security text to keep the text secret. It uses some mathematical computations called algorithms. Cryptography can be used to encrypt and decrypt data. The public key algorithms like RSA, DSA and ECC can be used for both encrypting and decrypting the data. ECC needs small size of keys compared to RSA.

The generation of keys in ECC is dependent on the elliptic curve given in Eq. 1.

$$y^2 = x^3 + ax + b \tag{1}$$

ECC uses scalar multiplication, in which point addition along with point doubling is included. The points on the elliptic curve are considered for exchanging the information [5]. For data encryption using ECC, the data is converted to points using scalar multiplication and the private key. The curve is sent to the receiver where it can be decrypted back to the plain text.

The work shows how the data (or plain text) is encrypted and decrypted using the points on the curve. The data (or plain text) is first converted to its ASCII values and then converted to the ASCII values into an integer. This integer is then converted to a point on the curve by using Eq. 1. After encryption, the reverse process is done to get the data (or plain text) back, which is called decryption.

2 Problem Statement

The method shows how the data is transmitted over IoT in a secured way by using ECC. In ECC to encrypt and decrypt the text, many researchers have been following different ways like a table containing conversion of characters, points on the curve or using by ASCII values. The above methods may not be fruitful to hide the message. So, a new technique is introduced in the paper to exchange the data in a more secured way [6].

In the new technique, there is no exchange of any kind of table with set of characters or the data is completely converted to ASCII values for transferring the data. The technique used here encrypts the data using the sender's private key. It is decrypted at the receiver's side by using the sender's public key.

3 Proposed Method

3.1 Flow Diagrams

Figure 1 shows how the data is transferred over IoT (the wireless network) and is converted to an elliptic curve. The data or plain text is given as input which is first converted to ASCII values and the ASCII values are grouped into sizes. Each group is padded with zeros of 10 bits at the end to avoid overlapping of the values. Now, each group is converted to an integer value. The integer value is used for converting to cipher text, which is like a point on the curve.

In the proposed method, the data is encrypted using private key of the sender. To encrypt and decrypt the data, Eq. 1 is used on both the sides.
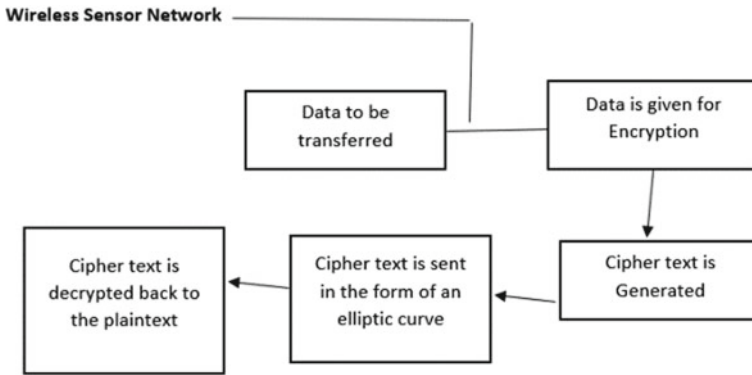
Fig. 1 Flow of the process

3.2 Process of Encryption and Decryption

The data is first converted to the to its corresponding ASCII values. The ASCII values are grouped into equal sizes (in Fig. 2). The size of the group is calculated by using

$$\text{groupsize} = \text{Length}\big[\text{InterDigits}\,(p,\ 65536)\big] + 1$$

The data is converted to ASCII values and are grouped according to the group size. A list is formed with the group size. To avoid overlapping of groups pad, each group by adding 10 bits of zeros at the end. A random x value is selected and a generator

Fig. 2 Steps in encryption process

'g', the range of x is between 1 and $p-1$. xg and xM_b are calculated using point multiplication [7]. M is a point on curve used in point multiplication operation. The point multiplication is computed using repeated addition. $Mm + kM_b$ is calculated using point addition along with point doubling. If M and N are points on the elliptic curve, the point addition is calculated as
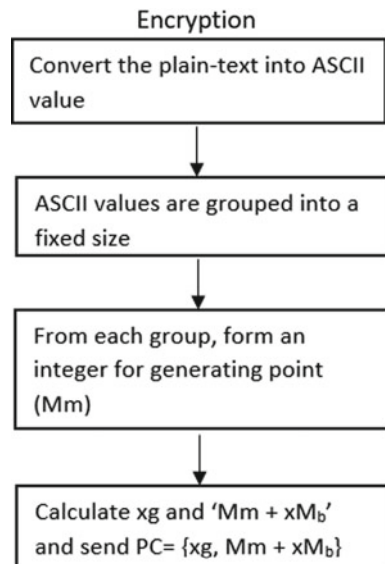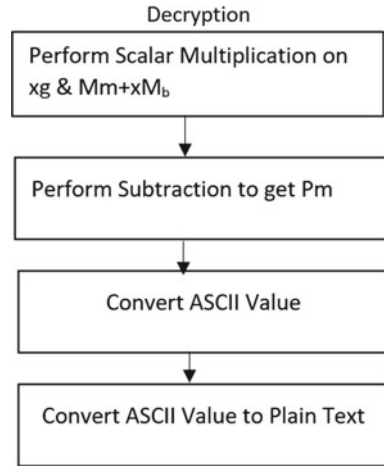
$$M + N = O(x_3, y_3)$$

The calculation xM is given by $xM = M + M + M + \ldots\ldots + x$ times. The point doubling operation for a point M on the curve is $O = 2\,M$. The cipher text $= \{xg, Mm + xM_b\}$ is sent to the receiver. Decryption is the reverse process of encryption (in Fig. 3). The cipher text is converted to plain text in decryption. The receiver has to decrypt the cipher text back to the plain text, which is the original message. Decryption uses both the decryption algorithm and a key to decrypt the message [8]. From the cipher text, left half (xg) is taken and point multiplication is performed by nB. Then $Mm + xM_b$ is subtracted by $nBxg$ which gives Mm. Convert Mm to ASCII. Then convert the ASCII values back to plain text.

3.3 Algorithms

Encryption Algorithm

1. Consider p as plaintext
2. Convert message p into ASCII
3. Set groupsize $=$ Length [IntegerDigits [p, 65,536]] $+ 1$, where $p =$ given integer i.e. ASCII values of Integer

Table 1 Comparison of time taken for encryption

No. of words	New technique (Time in seconds)	Implementation of text encryption using ECC (Time in seconds)
5	0.0012342	0.089
4	0.0010323	0.081
3	0.0008579	0.074
2	0.0007554	0.068
1	0.0006453	0.062

4. Then ASCII values are grouped into partitions
5. Pad the last 10 bits to zeros in each group
6. Select a random value x from $1-p-1$ and compute xg with point multiplication
7. Compute $c = + x$ using point doubling or addition
8. Calculate $Mm + xM_b$
9. Send the cipher text $\{xg, Mm + xM_b\}$.

Decryption algorithm

1. After getting the cipher text.
2. From the cipher text left half (xg) is taken and point multiplication is performed by nB.
3. Then $Mm + xM_b$ is subtracted by $nBxg$ which gives Mm.
4. Convert Mm to ASCII values.
5. Convert the ASCII values back to plain text.

4 Results and Performance Analysis

In any work, methods implemented must be compared with the previous implemented ones to know the best one. In the paper, the time taken for the method used is compared to the text encryption. The new technique takes less time compared to text encryption for encryption (Table 1).

The method implements decryption also. When compared, the new technique takes less time compared to text encryption for decryption (Table 2).

The data transferred using IoT must be secured. Data in any device must be securely exchanged and also the time taken for encryption or decryption must be less. So, with in less amount of time, the data is transferred securely.

5 Conclusion

The new process provides a high level of data protection to protect the data while transmitting in the IoT. The proposed method provides better security. As it provides

Table 2 Comparison of time taken for decryption

No. of words	New technique (Time in seconds)	Implementation of text encryption using ECC (Time in seconds)
5	0.00201344	0.103
4	0.00153789	0.102
3	0.00117832	0.098
2	0.00056782	0.088
1	0.00045633	0.081

more efficiency and confidentiality is provided by converting text to points on the curve. Even large amount of data is safely transferred over the IoT network. This paper shows how the data is encrypted using elliptic curve cryptography and to transfer the data securely. The method can also be used for large size of data.

References

1. Weber RH (2010) Internet of things—new security and privacy challenges. Comput Law Secur Rev 26(1):23–30
2. Ukil A, Sen J, Koilakonda S (2011) Embedded security for internet of things. In: Proceedings 2nd national conference on emerging trends and applications in computer science (NCETACS), Mar 2011, pp 1–6
3. Daniels W et al (2017) SμV-the security microvisor: a virtualisation-based security middleware for the internet of things. In: Proceedings ACM 18th ACM/IFIP/USENIX middleware conference industrial track, Dec 2017, pp 36–42
4. Sun H, Wang X, Buyya R, Su J (2017) CloudEyes: cloud-based malware detection with reversible sketch for resource-constrained internet of things (IoT) devices. Softw Pract Exp 47(3):421–441
5. Banerjee U, Juvekar C, Fuller SH, Chandrakasan AP (2017) eeDTLS: energy-efficient datagram transport layer security for the internet of things. In: Proceedings IEEE global communication conference, Dec 2017, pp 1–6
6. Manogaran G, Thota C, Lopez D, Sundarasekar R (2017) Big data security intelligence for healthcare industry 4.0. In: Cybersecurity for industry 4.0. Springer, Cham, Switzerland, pp 103–126
7. Yang Y, Liu X, Deng RH (2017) Lightweight break-glass access control system for healthcare internet-of-things. IEEE Trans Ind Inform 14(8):3610–3617
8. Ahmed S, Zaman A, Zhang Z, Alam KMR, Morimoto Y (2019) Semiorder preserving encryption technique for numeric database. Int J Netw Comput 9(1):111–129
9. Vucinic M, Tourancheau B, Rousseau F, Duda A, Damon L, Guizzetti R (2015) OSCAR: object security architecture for the internet of things. Ad Hoc Netw 32:3–16
10. Chervyakov N, Babenko M, Tchernykh A, Kucherov N, Miranda-Lopez V, Cortes-Mendoza JM (2019) AR-RRNS: configurable reliable distributed data storage systems for internet of things to ensure security. Future Gener Comput Syst 92:1080–1092
11. Raza S, Shafagh H, Hewage K, Hummen R, Voigt T (2013) Lithe: light weight secure CoAP for the internet of things. IEEE Sens J 1(10):3711–3720
12. Davoli L, Veltri L, Ferrari G, Amadei U (2019) Internet of things on power line communications: an experimental performance analysis. In: Smart grids and their communication systems. Springer, Singapore, pp 465–498

13. Debnath S, Nunsanga MV, Bhuyan B (2019) Study and scope of signcryption for cloud data access control. In: Advances in computer, communication and control. Springer, Singapore, pp 113–126
14. Boyd C, Hale B, Mjolsnes SF, Stebila D (2016) From stateless to stateful: generic authentication and authenticated encryption constructions with application to TLS. In: Proceedings of the cryptographers' track at the RSA conference. Springer, Cham, Switzerland, pp 55–71
15. Tyagi M, Manoria M, Mishra B (2019) A framework for data storage security with efficient computing in cloud. In: Proceedings of the international conference on advanced computing networking and informatics. Springer, pp 109–116
16. Louw J, Niezen G, Ramotsoela TD, Abu-Mahfouz AM (2016) A key distribution scheme using elliptic curve cryptography in wireless sensor networks. In: Proceedings of the IEEE 14th international conference on industrial informatics (INDIN), Jul 2016, pp 1166–1170
17. Kanda G, Antwi AO, Ryoo K (2018) Hardware architecture design of AES cryptosystem with 163-bit elliptic curve. In: Advanced multimedia and ubiquitous engineering. Springer, Singapore, pp 423–429
18. Dawahdeh ZE, Yaakob SN, Othman RRB (2016) A new modification for menezes-vanstone elliptic curve cryptosystem. J Theor Appl Inf Technol 85(3):290
19. Ferretti L, Marchetti M, Colajanni M (2019) Fog-based secure communications for low-power IoT devices. ACM Trans Internet Technol 19(2):27
20. Albalas F, Al-Soud M, Almomani O, Almomani A (2018) Security-aware CoAP application layer protocol for the Internet of things using elliptic curve cryptography. Power (mw) 15(3A):151
21. Khan S, Khan R (2018) Elgamal elliptic curve based secure communication architecture for microgrids. Energies 11(4):759

Sign Language Interpreter

Sanjay Kumar Suman, Himanshu Shekhar, Chandra Bhushan Mahto, D. Gururaj, L. Bhagyalakshmi, and P. Santosh Kumar Patra

Abstract This article addresses a design of an apposite system which provides a supportive hand for hearing and speaking challenged person to expediently communicate with normal people. Normally, a sign language is adopted by them for their communication which needs an interpreter to convert into user's understandable language. The proposed system is used for converting the sign language into voice and text and vice versa. The idea of the proposed project is to come up with a device that captures the gestures and converts it to voice output as well as in text output and also to capture the voice by speech recognition module and convert it to corresponding sign language by displaying on a screen with the help of various elements like microphone, camera, sign language database and display unit. For the general-purpose indoor implementation, a facial expression recognition system can also be additionally included.

Keywords Sign language · Interpreter · Speech recognition · Communicate · Gesture recognition · Facial expression recognition

1 Introduction

Earlier, it was very difficult for the deaf/dumb to communicate with a normal person because of the lack of a proper sign language and ease of understanding. But after the advent of sign language, the deaf/dumb now, are able to communicate not only with similarly placed, but also with normal people. At times, it is difficult to communicate

S. K. Suman · P. S. K. Patra
St. Martin's Engineering College, Secunderabad, Telangana, India

H. Shekhar
Hindustan Institute of Technology and Science, Chennai, TN, India

C. B. Mahto
Department Electrical Engineering, MIT Muzaffarpur, Muzaffarpur, Bihar, India

D. Gururaj · L. Bhagyalakshmi (✉)
Department of ECE, Rajalakshmi Engineering College, Chennai, TN, India
e-mail: Prof.Dr.L.Bhagyalakshmi@gmail.com

with normal people since, and it is not necessary that all the people whom they come across is acquainted with the sign language to understand what the deaf/dumb has to say. This is called as unintentional misunderstanding [1, 2]. In such cases, they have to hire an interpreter who can interpret their sign language and convert into speech for normal person to understand and vice versa. Still, there are some fallacies occurring in sign interpretation, predominantly in the field of business and transactions. To overcome this, we have an electronic interpretation device to stand by the deaf/dumb, so that they can communicate with ease. This would go a long way in establishing effective and reliable communication with the deaf/dumb and normal person without having to approach an interpreter.

Any movement of the hand or change in face that expresses a thought, emotion, feeling or reaction can be defined as a gesture such as: eyebrows movement and raising soldier are normal behavior used in our daily life. Sign language is a systematic and defined communication method in which each word or letter is assigned a specific gesture [3]. Here, a motion capture system is used for sign language conversion and a speech recognition system is used for speech conversion [4].

This idea can be executed using two different implementations, namely indoor and outdoor. Indoor module consists of facial expression recognition system. The only major issue will be collecting the list of all the words with their sign language. Creating the database is the most difficult task. Since there are many ways to interpret sign language, different possibilities can be used to design the system.

Facial expression recognition or emotion detection systems include three steps: face detection, feature extraction and facial expression recognition [5, 6]. The face detection algorithm for this system is based on the work of Viola and Jones. They proposed a face detection framework that can process images very fast and achieve high detection rates [7]. The database used is the most comprehensively tested Cohn Kanade database for a comparative study of the facial expression and emotion database [8]. Also, a local binary pattern is used for analyzing attitude emotions and textures [9, 10]. In addition, Microsoft's Kinect sensor XBOX 360 includes motion capture technology that can convert signatures to voice, and the camera decides to use it for scene capture.

2 Earlier and Current Issues

Earlier project of text-to-sign language conversion was limited the output to the PC base module and no portability [8]. A fact to be known is that sign language interpreters have cognitive abilities, perceptual skills and other characteristics that make them unique from others [11]. Further, there are about 12.3 million people having moderate to complete hearing loss in India, but we have only around 25,000 interpreters. Only a part of these deaf people (about 4.5 million) would not be able to succeed in a school for hearing people, whereas they can obtain education in special schools for deaf. They would then be introduced to sign language which might become part of life of the deaf and dumb community. A statistical analysis

reveals that around 478 govt. running school and 372 private schools receive fund from govt. agency for development of deaf children in India using oral approach rather than sign language [12].

Worse situation found in rural part of developing country is where the CODA does not receive an education due the distance. They have to struggle to attend the school and also to the need to work at home. Many deaf and mute people are talented at many fields in spite of their disability to speak or hear. The various fields include business, biology, psychology, arts, science, mathematics and computer, etc. Presently, there is a need for qualified interpreters in medical fields, businesses and offices for making the language translation easier.

Children of deaf adults, called as CODAs, often serve as interpreters in most parts of the world. However, in many developing countries, CODAs are either not qualified or reluctant to work as interpreters due to their inevitability. Another issue is dependency of the deaf parents on their parent or on relative, for nurturing care and education of their children. In this case, the children do not get proper education and even fundamental requirements. Also, many CODAs do not admit that they possess deaf parents due to the fear of discrimination and uncertain problems that might arise upon revelation [12].

3 Proposed Work

This paper describes two implementations namely indoor and outdoor. The indoor module contains additional features such as facial expression recognition and lip reading (optional) for more accurate results along with camera, microphone, speakers and display unit. By introducing this device for sign language interpretation, we can overcome the discrimination and difficulties that a deaf or mute person faces in the society while communicating with a hearing person. The deaf and mute community calls a normal person as "hearing person." The device would consist of a database of sign language visuals like animation [13] and the corresponding word display. The intermediate module would be the converter depending upon the process that is to be carried out.

The database of the sign language can be bifurcated to be used by different kinds of people, like for kids, smart class and schools for learning purposes with the help of smart screens and projectors. They can get educated along with normal children in normal schools and avoid going to special schools where people only of similar kind are enrolled. A universal language for example English can be used at an initial level. Depending upon the usage and purpose of application, it can be developed in other languages for better understanding. Regular updates by means of apps can be launched for updated and modified version of the sign language conversion tool.

4 System Model

This section presents the system model which comprises: voice input to sign input, sign input to voice input and facial expression recognition.

- **Module I**

  The first module, as shown in Fig. 1, consists of a speech recognition device, the database containing the video content to various actions in sign language, the matching device that converts the corresponding audio to respective video (sign language) and the output display unit. The sign language animations and symbols are loaded to the database of the device. When the mic captures the audio of a normal person, it gets converted to text internally and that will be matched with the corresponding interpreted sign output in the database. This converted output, i.e., the matched output, will be displayed on the display screen. An adaptive noise canceling microphone system is used here to capture the voice.

Speech Recognition

We can either make use of the design of Kaldi which is a free open-source toolkit for the purpose of speech recognition [14, 15]. Kaldi gives us a speech recognition system which is based on finite-state transducers (using the freely available OpenFst), together along with the detailed documentation and scripts for creating a complete recognition system. The only issue is that it gives only a considerable level of product satisfaction. Another idea is to use google speech recognition system which is comparatively quick and also easy which makes use of the IOT concept. There are various approaches for speech recognition as follows:

- *Template*: An unknown speech is compared with a set of pre-recorded words and alphabets (templates) to find the best match.
- *Knowledge*: A robust knowledge about variations in speech is hand coded into a system so that recognition is facilitated.
- *Statistical*: This is the method by which variations in speech are modeled statistically, using automatic, statistical learning procedure, typically the hidden Markov models, or HMM. This method is usually tedious and not up to the level of satisfaction.

Fig. 1 Speech-to-sign conversion module
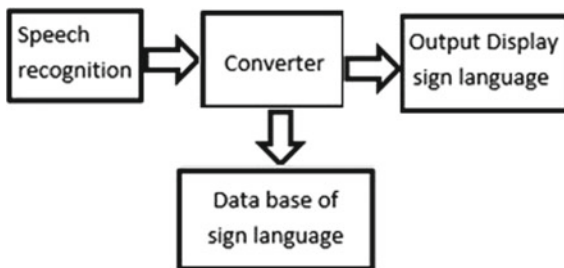
- *Learning*: Machine learning methods could be introduced such as neural networks and genetic algorithm/programming in order to overcome the disadvantage of the HMMs.
- *Artificial Intelligence*: The artificial intelligence approach attempts to mechanize the recognition procedure according to the way a person is applying his/her intelligence in visualizing, analyzing and finally making a decision on the *measured* acoustic features or data.
- *CMU Sphinx*: CMU Sphinx, is also called as Sphinx in short, is the general name of speech recognition systems which are developed at Carnegie Mellon University. There are three speech recognizers from Sphinx 2 to 4 and an acoustic model trainer which is Sphinx Train. In this project, Sphinx 4 can be used. It purely depends on the quality of output required, that we make the choice of the appropriate speech recognition tool. There are various sub-modules of the Sphinx.

Database

Words for speech recognition, images and motions (videos) of sign language all together create the database for the product.

Display Unit

The display unit is an ordinary screen like led display or a normal smart device like smart phone or iPad. The output this module will be displayed as animated video or still images demonstrated in Fig. 2 and Fig. 3, respectively.

Fig. 2 Animated sign language output

Fig. 3 Animated sign converted output

- **Module II**

The second module, shown in Fig. 4, the same device consists of depth camera, the database of speech audios matched to a corresponding sign or gesture of hands or body and finally an output speaker. The IR depth camera and its associated gesture recognition camera are depicted in Fig. 5 and Fig. 6, respectively. The sign gestures of the dumb person is captured by the depth camera and matched with the available database of voice audios and when the match to the action is found, the respective audio is played in the speakers.

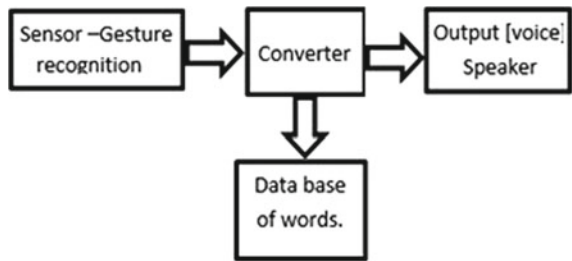**Fig. 4** Sign-to-speech conversion module
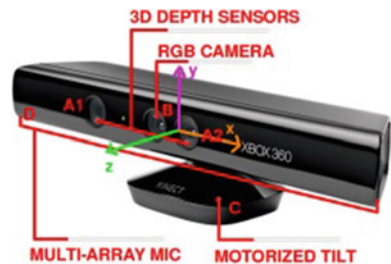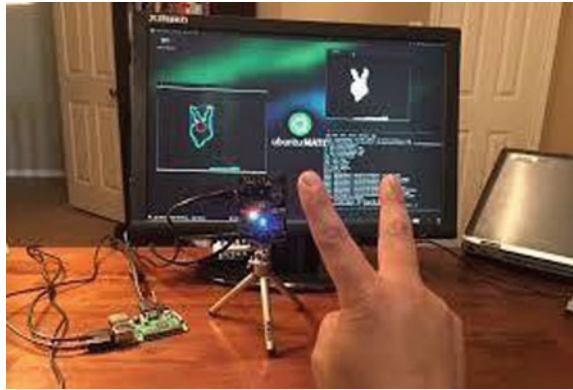

Fig. 5 IR depth camera

Fig. 6 Gesture recognition using camera

Camera

The Microsoft Kinect Sensor XBOX 360 was chosen to capture the technical and motion capture capabilities of converting signals to voice. Google Speech Recognition is used to convert speech into signatures. For android-based programs, only Google voice recognition is available.

Finally, you can combine the two components in Java by choosing the speech recognition program CMU Sphinx. The converter can also be configured and written in Java. Finally, a Java-based program can be written that are capable of speech recognition and motion capture. One can use this program to convert the two to each other. As a result, hearing-impaired people can easily talk to ordinary people in sign language in front of a suitable camera, and people behind the screen can easily understand it even if they cannot sign language. The reverse is also true.

Microsoft Kinect XBOX 360TM was released by Microsoft with various sensors within. There are three sensors such as depth, audio and RGB as shown in Fig. 5. The various sensors are engaged to detect movements and recognize bodily gesture and sound. This is also widely used in robotics and action recognition for creative designing in games [3].

Features

Figure 6 illustrates the gesture recognition using camera. The features of IR depth camera are divided into four parts:

- Part A is also called a depth sensor or 3D sensor. The combination of an infrared laser projector and CMOS sensor allows the Kinect sensor to process 3D scenes in ambient lighting conditions. Using infrared light from the projector in the area of consideration, the sensor receives reflections from various objects in the scene. The depth map correctly specifies the distance between the object's surface and the point visible to the camera. This is called time-of-flight because it sets up the depth map of the scene, taking into account the amount of time it takes light

reflected off an object from the sensor view to return to the light source. The optimal depth range for the sensor is 1.2–2.5 m.

- Part B is a 32-bit and high-resolution RGB camera. It has the ability to create a two-dimensional color video of a scene.
- Part C is called motor tilt, which is primarily related to the field of view.
- Part D includes a microphone. It is on a horizontal bar. This is the 4 microphone array. It is useful for environmental noise suppression, correct speech recognition and echo cancelation.

Voice Output

Miniaturized speakers are used so that it becomes handy and portable for the person and gives a level of comfort. These speakers are manufactured in smaller sizes than a normal loud speaker yet providing louder output tone.

- **Module III**

The module 3, shown in Fig. 7 comprises of facial expression recognition system. The interaction between human and computer can be made effective if the computer can recognize the emotional state of a human being. Information regarding a person's emotion is expressed in terms of facial expression. Hence, recognizing the facial expressions will let us know something about the emotional state of the person. However, it is hard to categorize facial expressions from normal images. In this problem neural network may be suitable because it can be used to improve its performance. Moreover, it is not necessary to know much about the features of the facial expressions to build the system.

An image containing a human face with an expression in the size of 96×72 pixels can be used as the input to the system. On an average there are 6 outputs representing each facial expression (with further advancements more expressions can be loaded depending upon usage). Each number represents a facial expression (smile, angry,

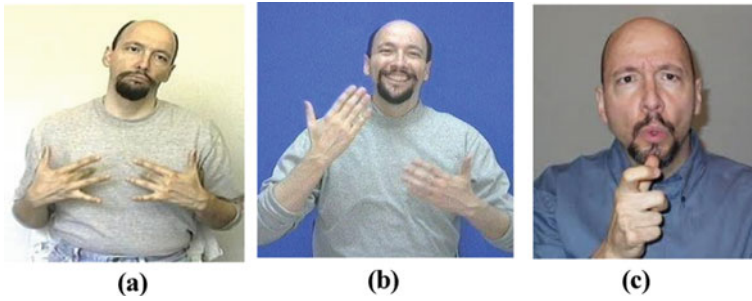**Fig. 7** Facial expression recognition module
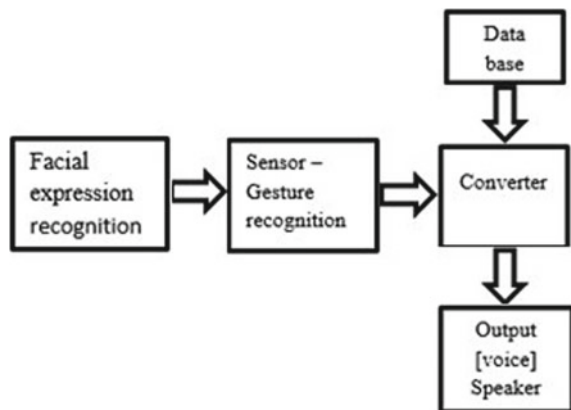
Fig. 8 The facial expression for: **a** sadness, **b** happiness and **c** reaction "who?"

fear, disgust, sadness, surprise). If that facial expression is present in the memory, then the number is 1 (one), and if not, it is 0 (zero) [6].

Facial Expression Recognition

A few examples of various facial expressions are explained as follows:

- *Sadness*: In Fig. 8a, this expression is to slightly lower the corners of the lips while raising the inner eyebrows. Darwin described this expression with a look that did not want to cry. The lower lip is lowered because the upper lip control is larger than the lower lip control. When a person cries and screams, close their eyes to protect them from the build-up of blood pressure in their face. So when I have the urge to cry and want to stop, I try to raise my eyebrows without closing my eyes.
- *Happiness*: This expression usually involves a smile: both corner of the mouth rising, shown in Fig. 8b.
- *A question such as-Who?*: This expression, Fig. 8c, is obtained as a result of inverted v-shaped eyebrows and curvy mouth with hand in the shape as shown in the picture

5 Implementation

(1) *Implementation 1*

The first method is to make a handy device like a wearable chain where the locket consists of the miniaturized camera, microphone, speaker and a memory device for the database. The output unit will be a display unit in the form screen either a smart phone or an iPad or a unique special purpose display screen. This display unit is also miniaturized one for portability and easy handling. The person wears the chain and activates the device, the device starts sensing the gesture made by the deaf or dumb person's hand (in front of the camera), and the gesture recognition system comes to work and translates the sign to corresponding voice output. Thus, this establishes a conversation between the deaf or mute person and a hearing person so that the sign is translated to voice.

Now, if the deaf or mute person has to understand the normal person, he/she activated the device for speech recognition.

(2) *Implementation 2*

The second method is for indoor purpose where the device along with facial expression recognition system is installed at a particular location inside the room (e.g., classroom). This method is exclusively for the classroom purpose.

6 Conclusion

This article presented the interpreter which can be used in a closed room or outside. Also, this device can be used for smart classes, library and public utility services like airport, bus stations, railway stations, hospitals, Internet hubs, hotels, restaurants, malls and offices. This could be more beneficial for communication at schools for the deaf/dumb so that they can feel themselves on par with normal person. Deaf and mute children are prone to be looked down upon by normal children, thereby creating an inferiority complex among themselves. This can be avoided by using the interpretational device which will remove the barrier of emotional differences between them and a normal child. Even though they lack the power to hear and speak, they are multiskilled personalities and excel in their own interests. Their potential and capability can be discovered to achieve greater heights in life.

References

1. Read MK (1977) Linguistic theory and the problem of Mutism: the contributions of Juan Pablo Bonet and Lorenzo Hervas Y Panduro. Historiographia linguistica 4(3):303–318. https://doi.org/10.1075/hl.4.3.03rea
2. Harvey MA (2003) Shielding yourself from the perils of empathy: the case of sign language interpreters. J Deaf Stud Deaf Educ 8(2):207–213. https://doi.org/10.1093/deafed/eng004
3. Arsan T, Ülgen O (2015) Sign language converter. Int J Comput Sci Eng Surv 6(4):39–51. https://doi.org/10.5121/ijcses.2015.6403
4. Kalsh A, Garewal NS (2013) Sign language recognition system. Int J Comput Eng Res 03(6). http://www.ijceronline.com/papers/Vol3_issue6/part%201/D0361015021.pdf
5. Piatkowska E (2010) Facial expression recognition system. Master thesis: technical reports. https://via.library.depaul.edu/cgi/viewcontent.cgi?article=1017&context=tr
6. Do C-T, Pastor D, Goalic A (2010) On the recognition of cochlear implant-like spectrally reduced speech with MFCC and HMM-based ASR. IEEE Trans Audio Speech Lang Process 18(5):1065–1068. https://doi.org/10.1109/TASL.2009.2032945
7. Viola P, Jones MJ (2004) Robust real-time object detection. Int J Comput Vision 57(2):137–154. https://doi.org/10.1023/B:VISI.0000013087.49260.fb
8. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture with local binary patterns. IEEE Trans Pattern Anal Mach Intell 7(7):971–987. https://doi.org/10.1109/TPAMI.2002.1017623
9. Wallhoff F, Schuller B, Hawellek M, Rigoll G (2006) Efficient recognition of authentic dynamic facial expressions on the feedtum database. In: IEEE international conference on multimedia and expo. https://doi.org/10.1109/ICME.2006.262433

10. Kanade T, Cohn JF, Tian Y (2000) Comprehensive database for facial expression analysis. In: Proceedings of the 4th IEEE international conference on automatic face and gesture recognition, Grenoble, France, pp 46–53. https://doi.org/10.1109/AFGR.2000.840611
11. Seal BC (2015) Psychological testing of sign language interpreters. J Deaf Stud Deaf Educ 9(1):39–52. https://doi.org/10.1093/deafed/enh010
12. Sugandhi PK, Kaur S (2021) Indian sign language generation system. IEEE Mag Comput 54(3):37–46. https://doi.org/10.1109/MC.2020.2992237
13. Halawani SM (2008) Arabic sign language translation system on mobile devices. Int J Comput Sci Netw Secur 8(1):251–256 (King Abdulaziz University, Jeddah, Saudi Arabia). https://www.kau.edu.sa/Files/830/Researches/56041_26352.pdf
14. Povey D et al (2011) The Kaldi speech recognition toolkit. In: IEEE 2011 workshop on automatic speech recognition and understanding, Hilton Waikoloa Village, Big Island, Hawaii, US. https://www.danielpovey.com/files/2011_asru_kaldi.pdf
15. Oliveira T, Escudeiro N, Escudeiro P, Rocha E, Barbosa FM (2019) The virtual sign channel for the communication between deaf and hearing users. IEEE revista iberoamericana de tecnologias del aprendizaje 14(4):188–195. https://doi.org/10.1109/RITA.2019.2952270

Noise Removal Filtering Methods for Mammogram Breast Images

Mudrakola Swapna and Nagaratna Hegde

Abstract Breast cancer detection in the early stage is an important factor to reduce the mortality rate. Mammogram examination is one of the best optimistic from various approaches used in the early detection of breast cancer at a different stage of cancer and the raw mammogram images are required to pre-process for better radiologist perception and to obtain an enhanced and clear image. It also helps to extract the Region of Interest from the processed image by using statistical feature methods to find the size and shape of the tumor. This paper is on an experimental study performed on sample mammogram images and applies different noise smoothing methods. Methods used to remove noise from the images by applying filtering methods like Gaussian Filter, Tri-State Filter, Mean Filter, Mean-Median Filter, Threshold Filter, Bilateral Filter, Wiener Filter, and Adaptive filter. The processed and obtained quality image will help doctors and radiologists to give an accurate impression on a patient case study. **Results**: quality of the image obtained on sample mammogram images of CBIS-DDSM dataset achieved min 80% of quality PSNR values.

Keywords Invasive · Breast cancer · TNM · Mammogram · Tumors · Gaussian · Mean · Median · Bilateral · Wiener and adaptive filter

1 Introduction

Breast cancer is the most common cancer type in India, which has taken lead over cervical cancer in comparison with past decades [1]. The anatomy of the breast has chest muscles, glandular tissues, and lobes. Breast lies on the pectoral muscles, glandular tissues help to produce milk and they look like bulb shape and structure is

M. Swapna (✉)
Osmania University, Hyderabad, India
e-mail: swapna0801@gmail.com

N. Hegde
Department of CSE, Vasavi College of Engineering, Hyderabad, India
e-mail: nagaratnaph@staff.vce.ac.in

arranged like wheels with 15–20 spokes on the nipple. The size of the breast is the fattiness of tissue [2, 3].

Symptoms of breast cancer are may or may not notice. The noticeable symptoms like lumps found under the armpit, deformation/dislocation/inverted nipples. The constant pain and swelling are other clinical observations. The study of Pathophysiology on breast cancer states that cancer cells are generated abnormally by DNA ruptures, damage or destructive behavior of genetic mutation is also reason [2, 4].

The factors that influence breast cancer can be biological or habitats. The biological reasons can be gender, family history, histological risk, age, menarche under 12 years' age, planning children's lately, and menopause after 50 years of age. Habitats are human-developed daily habits that lead to ill health factors smoking tobacco, intake of high-fat food, excuse intake of alcohol, and exogenous harmon usage: premenopausal and post premenopausal in women. [2, 4].

Breast cancer is classified into two types as non-invasive and invasive breast cancer. Non-invasive cancer found in milk ducts and the lining of lobules is known as ductal carcinoma. This type of cancer will not spread to other parts of the body. Invasive cancer is found in the milk ducts and spread outside the breast to armpits. The rarely found cancers are Metaplastic, Tubular, Mucinous, papillary, and Medullar [5].

Hormone Receptor-Positive: either from the estrogen or progesterone receptor one of the cells is positive then, it is declared as HR + cancer. In case both receptors are detected as negative then it is confirmed to be hormone receptor-negative. HER2 Positive: HER2 means Hormone Epidermal Growth Factors Receptor2, which create abnormal number of copies of the gene more than protein required which Otherwise it is said to be HER2 negative. Triple-Negative: In this case estrogen receptor, progesterone receptor, and HER2 are negative, then it is said to be triple-negative. These types are said to be invasive breast cancer [5].

Tumor node metastasis (TNM) is a tumor staging system used to define the stages of breast cancer. Where tumor helps to identify the primary tumor exists, node defines the tumor spread to lymph nodes or not, and also Metastasis defines speeded to other parts of the body. Grades 0, 1, 2, 3, represents from zero to one, two or more are absence and level of cancer.

Stage 0: Non-Invasive or In-Site cancer exists only in ducts.

Stage IA: (T1-N0-M0)

Stage IIA: (T0-N1-M0)/(T1-N1-M0)/(T2-N0-M0)

Stage IIB: (T2-N1M0)/(T3-N0-Mo)

Stage IIIA: (T0-T1-T2)/(T3-N2-M0)/(T3-N1-M0)

Stage IIIB: (T4-N0-N1/N2-Mo)

Stage IIIC: (T0-N3-M0)

Stage IV: (Any T-Any N-M1).

These are the above stages defined [6].

Different breast cancer detection methods are available, but primary and regular breast self-examination tests need to be done for every woman to find early symptoms. MRI examination is used to detect the BRCA gene mutation analysis for any rupture in DNA structure, biomarkers analysis, urinary exposal examination, and ultrasound examination will help to find the lesions margin and boundaries. Computer tomography examination will help to detect the ab-normal edges and shapes. Mammogram examination is an efficient method because it has low radiation, highly available in the lab, and less cost [7].

Organization of Paper: The study is on different preprocessing methods that apply to grayscale mammogram images. In this paper, Sect. 1 gives comprehensive information on background knowledge of breast cancer. Section 2 will give a detailed literature study on breast cancer. Section 3 represents different filtering approaches. Section 4 is on a discussion on the quality measure of image. Section 5 is on the conclusion and future scope.

2 Literature Survey

Early-stage detection of breast cancer its importance and effectiveness have been studied and proven in Malaysia that 34.1% of cancers detected at the last stage. After awareness programs on breast self-examination (BSE), an overall 30% of cases were detected at stage-III instead of leading to stage IV. The cost of testing was neutral [8]. It has been proved that there is a possibility of reoccurrence of breast cancer in few cases. Regular breast self-examination (BSE) was highly recommended by a doctor [9, 10].

Gene expression profiling will perform multivariable analysis on Luminal A, Luminal B, HER-2, and Basal. This outcome observed that 10% of people have reoccurrence over five years [11]. Biomarker detection is not yet widely accepted in clinical practices. The key confirmation was dependent on an expert pathologist only. It will also help in identifying the unhealthy tissue by assessing 58 pairs of tissues. It is observed that p01, p13, and p35 pair probes will detect breast cancer with up to 89% accuracy, 82% of sensitivity, and 94% specificity [12].

Breast Biopsy is one of the complicated methods to diagnose lesions or tumors. There are different types of biopsy like core-needle biopsy, fine-needle biopsy, and vacuum-assisted breast biopsy is efficient to compare the other two methods. They are used to mammographic screening, sonographic, MRI, and physical tests above examinations will eliminate the surgical biopsy and helps to removes the undetermined lesions and masses from the breast to remove confusion in future diagnosis. VABB test will have some practical complications like even it is painful during the procedure but at tolerance level pain, bleeding but easy to use, etc., collection of more portions will help in obtaining the best results and low errors in sampling [13, 14].

Removal of salt and pepper and Gaussian noise can be done using different denoising methods, but there is always the threat of losing actual data but using the Global Unsymmetrical Trimmed Median filter (GUTM) method will be appropriate to de-noising image without loss of data. The matrix image has a processing window size of 3 * 3 matrix, to check the noise or not, if the noise was found on the window, replace its pixel with median value, else move the processing window to the next step. The results having a PSNR value is 52.31 and an MSE value is 0.37 [15].

Thresholding is one of the image processing methods. It can be implemented using fuzzy arithmetic; filter matching helps to detect the tiny particles on the image. There are various thresholding algorithms like histogram entropy, histogram shape, and clump of gray level data, domestically adaptive characteristics and spatial data are used to image transformation methods to investigate the specious [16, 17].

Breast Imaging is an efficient approach to detect breast cancer. It is not efficient at the early stage of breast cancer-detecting primary tumors with less than 1 mm in size is also very difficult and has an ambiguity in result decision. Imaging will not only help in detection but in treatment, detailing of staging but also to follow-up after or while treatment [18].

Breast **PET/CT** (Positron Emission Tomography/Computer Tomography) test is a high resolution scanning specially used for breast diagnosis and treatment evaluation [19]. PET has proved that 68% detection for small tumors and 92% detection for large tumors. Major limitation is difficulty in detection of tumors less in size [20]. Fully field digital mammogram (FFDM) images can use to develop 3D view, which helps in examination of CC view and MLO view. Breast images can be synthetic 2D, FFDM, or digital breast tomosynthesis (DBT) for breast analysis. Among above methods. DBT will have 25% more accuracy in detection rate [21, 22].

Breast PET/CT (Positron Emission Tomography/Computer Tomography) test is a high-resolution scanning specially used for breast diagnosis and treatment evaluation [19]. PET has proved that 68% detection for small tumors and 92% detection for large tumors. The major limitation is the difficulty in the detection of tumors less in size [20]. Fully field digital mammogram (FFDM) images can use to develop a 3D view of the breast. It helps in the examination of the CC view and MLO view. Breast images can be synthetic 2D, FFDM, or digital breast tomosynthesis (DBT) for breast analysis. Among the above methods, DBT will have 25% more accuracy in detection rate [21, 22].

Contrast Enhancement Mammogram Techniques: Detection of tumors on the dense breast is very abnormal. To overcome this problem, we can use contrast-enhanced spectral mammography (CESM) approach. To have clear visibility of a superficial vein, tumors, and calcification. The examination procedure is similar to a regular mammogram test, but the clarity of the image, Iodinated contrast media is injected with a dosage ratio of 1.5: 1 (ML: KG) weight of a person. Images are viewed in MLO and CC views, with the vision of a high-energy image. These images will help to classify the masses, size of masses, etc. [23, 24].

A comparative analysis on various filtering methods with their quality measure as done by different researchers has been shown in Table 1.

Table 1 Comparison study on various filtering methods on images

S. No.	Author details	Filter method	Noisy type	Formula	Advantages	Disadvantages
1	Deng and Cahill [25]	Gaussian filter	Gaussian noise, random noise	$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$	Noise reduction	Lose fine line details on image Salt and pepper noise can't handle
2	Chen et al. [26]	Tri-state filter	Salt and pepper noise	$X_{pq}^{TSM} = \begin{cases} Y_{pq} & T \geq a_1 \\ X_{pq}^{CWM} & a_2 \leq T < a_1 \\ X_{pq}^{SM} & T < a_2 \end{cases}$	Fine details of image can be preserve	Not efficient for salt and pepper
3	Mahmood et al. [27]	Median filter	Random noise salt and pepper noise	$y[m,n] = \text{median}\{x[i,j], (i,j) \in w\}$	Preserve sharp edges	False noisy edges may be created
4	Banerjee et al. [28]	Mean filter	Random noise	$\hat{f}(x,y) = \frac{1}{mn}\sum_{(s,t)\in s_{xy}} g(s,t)$	Noise reduction	Blur image are formed at edges
5	Sir et al. [16]	Threshold filter	Convert gray scale image to binary image	For black pixel $I_{i,j} < T$ For white pixel $I_{i,j} > T$	Quick in transmission	Boundary clearly will reduce
6	Zhang [29]	Bilateral filter	Random noise	$BF[I]_p = \frac{1}{u_p}\sum_{q\in S} G_{\sigma s}(\|P-Q\|)G_{\sigma r}(I_p - I_q)I_q$	Preserve edges	Speed is limited
7	Joseph et al. [30]	Wiener filter	Blurriness removal, additive noise	$\overline{F}(u,v) = \left[-H(u,v)/\left(\|H(u,v)\|^2\right) + K\right]G(u,v)$	Recover noisy and low resolution image	Cost of computation is very high in compare with smoothing technique

3   Preprocessing Mammogram Images Using Various Filtering Methods

Mammogram images are larger in size, approximately the pixel of 4000 × 5000 and size 100 μm. These images are complex to mobilize over the internet for remote access for virtual clinical trials and studies. Original images are compressed or resized and can be compressed into a maximum of seven times with lossless compression. Resize of images can be done using different algorithms like nearest neighboring and radial basis function interoptability, etc. [31, 32].

The quality of grayscale images is the most critical factor in the detection of disease. The noise images can be Gaussian noise, speckle noise, random, and salt and pepper noise. Noises can be removed by using various denoising methods. The quality of the image can be evaluated through Peak-Signal-Noise-Ratio, Mean Square Error. Gaussian noise is very difficult to remove noise, and it will try to remove the low and high pass signals. Salt and pepper noise can be removed efficiently using a modified decision-based algorithm [30, 33, 34].

The details of dataset used for our work is shown in Table 2.

3.1   Gaussian Filter

Filters are used to remove or reduce the noises on the images. A few of them have experimented included in this study. Gaussian Filter uses to remove the blur edges of the image by increasing the peak intensity at boundaries and also helps in the removal of noisy data. It is a low pass filter [35]. Gaussian filter was applied to 2D images. Where a sigma (standard deviation), the kernel window size is 3 * 3 matrix. It is a non-uniform low-pass filter. Gaussian filters will preserve the brightness of the image, kernel coefficients are dependent on values at the edge, the values are a mask to zero, *y* is the axis. Gaussian filters are used to generally for edge detection [36].

Table 2 Details of dataset used for experimental analysis

Details	Value
Name of the dataset name	The digital database for screening mammography
Authors	Michael Heath, Kevin Bowyer, Daniel Kopans, Richard Moore and W. Philip Kegelmeyer
Publisher	Medical Physics Publishing, 2001. ISBN 1-930,524-00-5
Event	Proceedings of the fifth international workshop on digital mammography
No of Cases	2620 available
No of cases studied	100 cases

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2 + y^2)}{2\sigma^2}} \quad {}^1 \tag{1}$$

3.2 Tri-State Filter

Salt and Pepper Noise can be used to eliminate by using Tri-State Filter. The size of the image is termed into 2D Matrix, the 3×3 Matrix and near elements must be either 0 or 255, any conflict in similarity then there exists noisy data and it can be trimmed from variant to the median value. It is one of the latest frameworks merged with approaches like center-weighted median and standard median filter (SMF) which are used to detect the impulse pixel. Identified pixels are modified with some fixed threshold value. They are verified by sliding the 3 * 3 matrixes over the entire image. It works with a center-weighted median (CWM) filter that uses the concept of assigning more weights to the central part of the image. SMF is a method used to eliminate the impulse pixel to remove noise and protect the edges of the content [26].

$$X_{pq}^{TSM} = \begin{cases} Y_{pq} & T \geq a_1 \\ X_{pq}^{CWM} & a_2 \leq T < a_1 \\ X_{pq}^{SM} & T < a_2 \end{cases} \tag{2}$$

where as $a_1 = |Y_{pq} - X_{pq}^{SM}|$, $a_2 = |Y_{pq} - X_{pq}^{CWM}|$ and T is defined biased threshold value.

3.3 Median Filter

Speckle noise is a degraded image cause due to radar signal fluctuations reflect occur at a time of examination. The above types of noise can be eliminated using mean filter, Wiener filter, adaptive median filter, and median filtering. The median filter is a more efficient method in comparison with other methods. Other methods may blur the edges and line feature of the image. Median values are calculated by sorting the numerical values. The calculated median pixel values are replaced with line and edge pixels and also isolates the noise pixel [27].

$$y[m, n] = \text{median}\{x[i, j], (i, j) \in w\} \tag{3}$$

where as 'w' is a near-by window $[m, n]$ is a central location axis.

3.4 Mean Filter

A mean filter method is used to smoothen images. To improve the quality of images, the mean method will help to find the intensive pixel value and they are replaced with the average value of neighboring pixels of the image. In filter will travel from one pixel to another pixel by replacing the average pixel value of each pixel. When we are replacing the edge pixel with the average value, the image may be blur then share edges can't be projected in view. The advantages of the mean filter are single noisy pixel will not impact the mean value [28].

$$\hat{f}(x,\ y) = \frac{1}{mn} \sum_{(s,\,t)\in s_{xy}} g(s,\ t) \Bigg\}$$ (4)

S_{xy} is a window size of sub image $m * n$ Are central point at (x, y) coordinates.

F is a restored image at point (x, y).

3.5 Threshold Filter

The process will transform a grayscale image into black-white images with a specific defined cut-off value or threshold value. There is lower and upper range value. This will help to turn a pixel into either white or black. The threshold values are defined automatically through an approach where it calculates a set of the 8-bit mean of the original image. It also helps to minimize the background noise. The images are split into foreground and background, below thresholds background and above threshold foreground. Finally, the average mean of new images is calculated, the difference values must be within the limit, then apply the change else change the threshold value and repeat the process from the initial steps [16].

Image intensity is defined as $I_{i,\,\mathrm{j}}$.

Threshold value is constant T.

$$\begin{aligned} \text{For black pixel} \quad I_{i,\,\mathrm{j}} &< T \\ \text{For white pixel} \quad I_{i,\,\mathrm{j}} &> T \end{aligned}$$ (5)

3.6 Bilateral Filter

A bilateral image is used to remove the noise from the image, which is a nonlinear and edge-preserving method. In this method, every pixel is altered with the average

weights of the surrounding pixels. It also considers the spatial nearby photometric range. The bilateral filter is calculated using the below formula [29].

$$\text{BF}[I]_{\text{p}} = \frac{1}{w_{\text{p}}} \sum_{q \in S} G_{\sigma s}(||P - Q||) G_{\sigma r}\big(I_{\text{p}} - I_{\text{q}}\big) I_{\text{q}} \tag{6}$$

W_{p} is a normalized factor

$$w_{\text{p}} = \sum_{q \in S} G_{\sigma s}(||P - Q||) G_{\sigma r}\big(I_{\text{p}} - I_{\text{q}}\big)$$

Overfitting is also another issue in filtering to moderate the filtering is used. Bilateral filtering also performs the iterative approach, incremental filtering to overcome from Gaussian blur problem [37].

3.7 Wiener Filter

The Wiener filter is one of the methods used to reduce the mean square error (MSE) and improve the quality of an image. This method is specially used to denoise the additive noise. It will try to identify the unknown signals and also eliminates the blurriness in the images. The Wiener filter has the drawback of losing the fine details of the image [27].

$$\overline{F}(u,\ v) = \big[{}^{-}H(u,\ v)\big/\big(|H(u,\ v)|^2\big) + K\big] G(u,\ v) \tag{7}$$

In which,

$F'(u,\ v) = $ The Estimate,

$G(u,\ v) = $ degraded image,

$H(u,\ v) = $ degradation image,

$H\big(u',\ v\big) = $ Complex$'$ Conjugate of $H\ (u,\ v)$,

K=constant.

The two-dimensional image is $F\ (u,\ v)$, it is inverted on high pass filter and for parameter K applies a low pass filter [38].

3.8 Adaptive Median Filter

This method is used to remove the noise from the images. It helps to reduce the distortion on boundary thickness and thinness and also removes impulse pixels. First, it will determine the pixel which is high impulse noise. Each and every pixel is compared with the nearby pixel and the diameter of the nearby pixel varies and has an adjustable size. Now, it will compare the similarity among the pixels and identify impulse pixels. The impulse pixels are replaced with a median pixel value of near-by pixel area and algorithm [23, 39].

Algorithm:

Step 1:

Xa = Zmed−Zmin

Xb = Zmed–Zmax

if Xa > 0 AND Xb < 0, go to step 2

else increase the window size

if window size < Smax, repeat step 1

else output Zxy

Step 2:

ya = Zxy−Zmin

yb = Zxy−Zmax

if ya > 0 AND yb < 0,

output Zxy

else output Zmed.

4 Results and Discussion

The quality of the images can be measure either through MSE and PSNR methods. In this study we applied different image filtering methods like Gaussian filter, tri-state filter, mean filter, mean-median filter, threshold filter, bilateral filter, Wiener Filter, and adaptive filter are used to remove noise from the image to improve the quality.

4.1 Mean-Square Error (MSE)

This method is used to calculate the average mean square of difference [28].

$$\text{MSE} = \frac{1}{n} \sum_{i=1} \left(Y_i - \hat{Y}_i \right)^2 \tag{8}$$

MSE=Mean Square Error

$n =$ Data Points

$Y_i =$ Observed values

$\hat{Y}_i =$ predicted values

4.2 Peak-Signal–Noise-Ratio (PSNR)

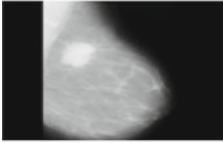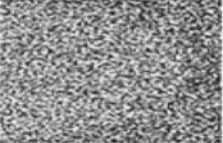This method overcome the drawback of MSE (factor of image intensity scaling), PSNR is measure in dB. S is defined as pixel value. Optimal value must PSNR = 30 dB [28].
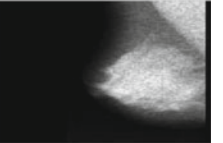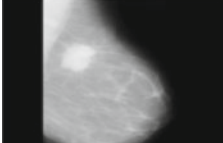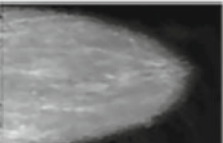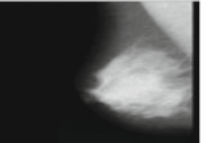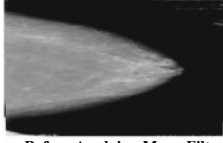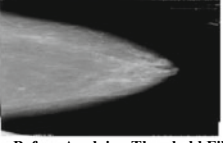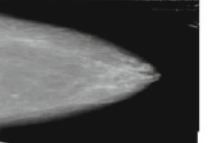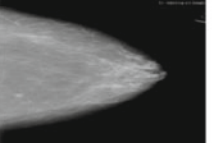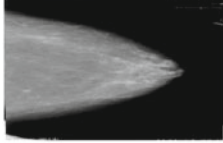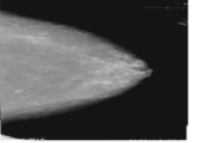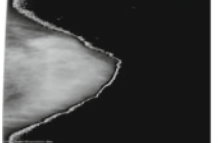
$$\text{PSNR} = -10 \log_{10} \left(\frac{e_{\text{MSE}}}{S^2} \right) \tag{9}$$

Various filtering methods were applied and the quality of images is measured which is shown in Tables 3 and 4.

Table 3 Results after applying various filtering methods on images

Filtering approach	MSE (Mean square error)			PSNR (Peak-signal-noise-ratio)		
	Image A	Image_B	Image_C	Image A	Image_B	Image_C
Gaussian filter	0.0271	0.0259	0.0263	63.8289	64.0210	63.9626
Tri-state filter	0.0259	0.0258	0.0266	64.1674	64.1754	64.0478
Median filter	0.0271	0.0259	0.0263	63.8289	64.0210	63.9626
Mean filter	0.0261	0.0249	0.0253	63.8279	64.0335	63.9345
Threshold filter	0.0252	0.0259	0.0250	64.1461	64.0276	64.1792
Bilateral filter	0.0259	0.0258	0.0266	64.0199	64.0339,	63.9145
Wiener filter	0.0252	0.0263	0.0272	64.8289	64.0210	63.8289
Adaptive filter	0.0251	0.0250	0.0258	64.1674	64.1754	64.0478

Table 4 Images obtained after applying various filtering methods

Noisy Blur Image	**Sample Salt and Pepper Noisy**	**Speckle Noise Image**
After Applying Gaussian Filter	**After Applying Tri-State Filter**	**After Applying Median Filter**
Before Applying Mean Filter	**Before Applying Threshold Filter**	**Before Applying Bilateral Filter**
After Applying Mean Filter	**After Applying Threshold Filter**	**After Applying Bilateral Filter**
Before Applying Wiener Filter	**Before Applying AMF**	
After Applying Wiener Filter	**After Applying AMF**	

5   Conclusion

Mammogram images are one of the optimum approaches for the detection of breast
cancer, but it is difficult to detection of small tumors as they merge with noise. This
study will help to remove different types of noise on images and also helps the
radiologist to perceive images in various shades and different perceptions to see the
faultiness in the image. After applying filtering methods, we can evaluate the quality

of the image using mean square error and peak Signal-to-noise ratio and calculate we use squared error and peak error, respectively. The value of MSE is lower, which means that lower the error rate. PSNR values are dependent on 8-bit image and 12-bit image values range from 30 to 50 dB and 60 to 80 dB, respectively. The MSE values are recorded lower for adaptive filtering. PSNR values are lower quality means that reconstructed images are very poor. Higher PSNR values are high means quality is good. PSNR values for bilateral filtering is a method that acquires maximum efficient values. The average accuracy of MSE is 0.0259 and the average accuracy of PSNR is 64.05%.

References

1. Rangarajan B, Shet T, Wadasadawala T, Nair NS, Sairam RM, Hingmire SS, Bajpai J (2016) Breast cancer: an overview of published Indian data. S Asian J Cancer 5(3):86
2. Alkabban FM, Ferguson T (2019) Cancer, breast. In: StatPearls (Internet). StatPearls Publishing
3. Fadhil SS, Dawood FAA (2021) Automatic pectoral muscles detection and removal in mammogram images. Iraqi J Sci 676–688
4. Mustafa M, Nornazirah A, Salih F, Illzam E, Suleiman M, Sharifa A (2016) Breast cancer: detection markers, prognosis, and prevention. IOSR J Dent Med Sci 15(08):73–80
5. Sharma GN, Dave R, Sanadya J, Sharma P, Sharma KK (2010) Various types and management of breast cancer: an overview. J Adv Pharm Technol Res 1(2):109
6. Hortobagyi GN, Edge SB, Giuliano A (2018) New and important changes in the TNM staging system for breast cancer. Am Soc Clin Oncol Educ Book 38:457–467
7. Panetta K, Samani A, Agaian S (2014) Choosing the optimal spatial domain measure of enhancement for mammogram images. Int J Biomed Imaging 2014
8. Soliman H, Abouelazayem M, Elkorety M, Nouh MA, Touny EM, Abdalla HM (2021) Impact of molecular profiling of breast cancer on the rate of locoregional recurrence in young versus old female patients. Cureus 13(1)
9. Moey SF, Mohamed NC, Lim BC (2021) A path analytic model of health beliefs on the behavioral adoption of breast self-examination. AIMS Public Health 8(1):15–31
10. Yeshitila YG, Kassa GM, Gebeyehu S, Memiah P, Desta M (2021) Breast self-examination practice and its determinants among women in Ethiopia: a systematic review and meta-analysis. PLoS ONE 16(1):e0245252
11. Hernandez LI, Araúzo-Bravo MJ, Gerovska D, Solaun RR, Machado I, Balian A, Botero J, Jiménez T, Zuriarrain Bergara O, Larburu Gurruchaga L, Urruticoechea A (2021) Discovery and proof-of-concept study of nuclease activity as a novel biomarker for breast cancer tumors. Cancers 13(2):276
12. Park HL, Hong J (2014) Vacuum-assisted breast biopsy for breast cancer. Gland Surg 3(2):120
13. Sennerstam RB, Franzén BS, Wiksell HO, Auer GU (2017) Core-needle biopsy of breast cancer is associated with a higher rate of distant metastases 5 to 15 years after diagnosis than FNA biopsy. Cancer Cytopathol 125(10):748–756
14. Piciu A, Piciu D, Polocoser N, Kovendi AA, Almasan I, Mester A, Morariu DS, Cainap C, Cainap SS (2021) Diagnostic performance of F18-FDG PET/CT in male breast cancers patients. Diagnostics 11(1):119
15. Jagadesh BN, Kumari LK (2021) A GLCM based feature extraction in mammogram images using machine learning algorithms. Int J Cur Res Rev 13(05):145
16. Sir K (2021) The impact of different image thresholding based mammogram image segmentation—a review. Glob J Comput Sci Technol
17. Zhang P, Li F (2014) A new adaptive weighted mean filter for removing salt-and-pepper noise. IEEE Signal Process Lett 21(10):1280–1283

18. Satoh Y, Kawamoto M, Kubota K, Murakami K, Hosono M, Senda M, Sasaki M, Momose T, Ito K, Okamura T, Oda K (2021) Clinical practice guidelines for high-resolution breast PET. Ann Nucl Med 1–9

19. Yang SK, Cho N, Moon WK (2007) The role of PET/CT for evaluating breast cancer. Korean J Radiol 8(5):429

20. Naeim RM, Marouf RA, Nasr MA, Abd El-Rahman ME (2021) Comparing the diagnostic efficacy of digital breast tomosynthesis with full-field digital mammography using BI-RADS scoring. Egypt J Radiol Nucl Med 52(1):1–13

21. Badal A, Sharma D, Graff CG, Zeng R, Badano A (2021) Mammography and breast tomosynthesis simulator for virtual clinical trials. Comput Phys Commun 261:107779

22. Bandyopadhyay SK (2010) Pre-processing of mammogram images. Int J Eng Sci Technol 2(11):6753–6758

23. Mehmood Gondal R, Lashari SA, Saare MA, Sari SA (2021) A hybrid de-noising method for mammogram images. Indonesian J Electr Eng Comput Sci 21(3):1435–1443

24. Anwar R, Farouk MA, Hamid WRA, El Maati AAA, Eissa H (2021) Breast cancer in dense breasts: comparative diagnostic merits of contrast-enhanced mammography and diffusion-weighted breast MRI. Egypt J Radiol Nucl Med 52(1):1–13

25. Deng G, Cahill LW (1993) An adaptive Gaussian filter for noise reduction and edge detection. In: 1993 IEEE conference record nuclear science symposium and medical imaging conference. IEEE, pp 1615–1619

26. Chen T, Ma KK, Chen LH (1999) Tri-state median filter for image denoising. IEEE Trans Image Process 8(12):1834–1838

27. Mahmood NH, Razif MR, Gany MT (2011) Comparison between median, unsharp and wiener filter and its effect on ultrasound stomach tissue image segmentation for pyloric stenosis. Int J Appl Sci Technol 1(5)

28. Banerjee S, Bandyopadhyay A, Bag R, Das A (2015) Sequentially combined mean-median filter for high density salt and pepper noise removal. In: 2015 IEEE international conference on research in computational intelligence and communication networks (ICRCICN). IEEE, pp 21–26

29. Zhang M (2009) Bilateral filter in image processing

30. Joseph AM, John MG, Dhas AS (2017) Mammogram image denoising filters: a comparative study. In: 2017 Conference on emerging devices and smart systems (ICEDSS). IEEE, pp 184–189

31. Yaffe MJ Digital mammography. Springer. http://eknygos.lsmuni.lt

32. Prasad P (2016) Color and gray scale image denoising using modified decision based unsymmetric trimmed median filter

33. Maheswari VU, Raju SV, Reddy KS (2019) Local directional weighted threshold patterns (LDWTP) for facial expression recognition. In: 2019 Fifth international conference on image information processing (ICIIP). IEEE

34. Maheswari VU, Prasad GV, Raju SV (2021) Facial expression analysis using local directional stigma mean patterns and convolutional neural networks. Int J Knowl-based Intell Eng Syst 25(1):119–128

35. Ramani R, Vanitha NS, Valarmathy S (2013) The pre-processing techniques for breast cancer detection in mammography images. Int J Image Graph Sign Process 5(5):47

36. Young IT, Van Vliet LJ (1995) Recursive implementation of the Gaussian filter. Signal Process 44(2):139–151

37. Hariraj V, Khairunizam W, Vikneswaran V, Ibrahim Z, Shahriman AB, Zuradzman MR, Rajendran T, Sathiyasheelan R (2018) Fuzzy multi-layer SVM classification of breast cancer mammogram images. Int J Mech Eng Tech 9(8):1281–1299

38. Safaei N, Smadi O, Safaei B, Masoud A (2021) A novel adaptive pixels segmentation algorithm for pavement crack detection

39. Hwang H, Haddad RA (1995) Adaptive median filters: new algorithms and results. IEEE Trans Image Process 4(4):499–502

Design and Implementation of Security Enhancement for Trusted Cloud Computing

Shiv Kumar Tiwari, Subhrendu G. Neogi, and Ashish Mishra

Abstract The concept of cloud computing is providing dynamic, scalable resources that are delivered over the Internet. Access to remote computer resources is made available to users through this service, and they only pay for the services that they actually use, at the time that they use them. However, for cloud users, the security of the information that is stored in the cloud is the most important concern that they have regarding cloud computing services. As a result of its ability to provide customers with on-demand, flexible, dependable, and low-cost services that are also scalable, cloud computing has experienced rapid growth in recent years. This invention relates to the design and implementation of an algorithm to improve cloud security with the goal of ensuring the security of information at the point of storage in the cloud. For a data security concept that emphasizes increased confidentiality and authenticity for cloud data at the cloud storage end, as well as experiment analysis to verify the approach's effectiveness and efficiency, a security strategy is established and developed.

Keywords Encryption · Security issues · Decryption · Confidentiality · Authentication · Cryptography · Cloud computing

1 Introduction

"Cloud computing" is an expression that refers to the evolution of a large number of computers that are networked, virtualized, and organized in a way that allows them to support portable workloads. The use of cloud computing is becoming increasingly popular. An application or service that runs on a distributed network and that is accessible through the use of common Internet networking protocols and networking standards is known as a distributed network service.

S. K. Tiwari (✉) · S. G. Neogi
Amity University, Gwalior, India
e-mail: sshiv.tiwari22@gmail.com

A. Mishra
GGITS, Jabalpur, India

$$Cloud = virtualization + abstraction$$

It shields users and developers from the specifics of system implementation, such as the fact that programs run on a physical system that is not specified, data is stored in an unknown place, and system administration is outsourced to a third party. For example, the AZURE platform, AMAZON services, and GOOGLE, among others. Virtualization is used in the cloud model to virtualized systems by pooling and sharing resources among multiple computers. Cloud computing is the calculation of a variety of resources and the delivery of those resources through a network of computers (Internet). Instead of preserving data on one's own computer or upgrading the application desires on one's own computer, it is possible to do so through a network (the Internet) [1, 2]. It enables users (individuals and organizations) to access software and hardware that is managed by third parties from a distance by allowing them to connect to the Internet. The term "cloud" refers to this form of network. Resources in the cloud can be expanded indefinitely, obtained at any moment, and utilized as needed. It dynamically distributes everything as a service on the Internet based on the demand of the user, including the operating system, the network, the hardware, the software, the resources, and the storage, among other things. The strengths and shortcomings of every computing paradigm are used to determine the degree to which it is accepted [3]. The cloud architecture is composed of three components: characteristics, delivery model, and deployment model. In addition to on-demand self-service, broad network access, resource pooling, quick elasticity, and measured service [1, 4], cloud computing has a number of other characteristics, including: The following are the cloud computing delivery methods [1, 4].

It is possible to subscribe to software through the use of the term "Software as a Service" (SaaS) in the context of the cloud. Gmail, Google Drive, and DropBox, for example, are all completely free services. When it comes to cloud computing, the term "Platform as a Service" (PaaS) refers to the model in which the cloud provides users with a platform or environment that they can use to run their applications over the Internet. Cloud-based services such as Google Gears and Microsoft Azure, for example, are both available. Cloud computing, also known as Infrastructure as a Service (IaaS), is a model in which a cloud service provider provides users with processing, storage, and networking capabilities on demand. A virtual representation of the infrastructure is provided to the user, but the actual physical infrastructure is handled by service providers located in remote locations. Cloud computing services such as Amazon Web Services and Google's Compute Engine, for example, are both available.

A private cloud is one that is owned and used by a specific organization that has complete control over the virtualized resources. A public cloud is a computer resource that is owned and provided for general public use by a specific organization or firm in order to provide access to computing resources at the lowest possible cost. The community cloud is a shared resource among many groups or businesses. Hybrid clouds are formed when more than two clouds combine to form a single cloud.

There are a number of difficulties [5] that arise in a cloud computing environment, including privacy, security, performance, load balancing, and reliability. The data

security issue is the most significant of these concerns [6, 7]. In order to improve data security at the cloud end, secure cloud architectures [8] have been proposed. Cryptography is the most effective method of protecting our information. Different encryption systems [9] for data security have been in use for many decades and are still in use today.

The paper is organized as follows: Sect. 2 discusses relevant work, Sect. 3 discusses the proposed technique, Sect. 4 discusses the results, and Sect. 5 discusses the conclusion.

2 Related Works

Security issues relating to cloud computing have received a great deal of attention in recent years. A number of approaches have been developed in order to achieve safe data storage at the cloud end. Table 1 illustrates the various existing research on cloud security, as well as the approaches that have been employed to address difficulties linked to the security of data.

The following algorithms, which are commonly used in cloud security, are briefly discussed in order to make our proposed work more accessible.

2.1 Modern Cryptography

Despite the advancement of encryption algorithms, such as RC6, AES, and DES, 3DES, and BlowFish, they continue to play an important role in data security in cloud computing [10–12]. The examination of those encryption methods that pass the randomness testing has been carried out in a cloud computing environment (Amazon EC2) utilizing NIST statistical testing [9].

Table 1 Recommended algorithms and layers

Security layer	Recommended algorithms
Confidentiality	Combination of RSA, ECC, RC6, Blowfish
Authentication	1. Username and password verification 2. Security question approval 3. Security token verification 4. One-time password verification 5. IP address and MAC address verification 6. Server verification using Kerberos
Access control	ABAC, RBAC, combination of RBA of ABAC and RBAC

2.2 Searchable Encryption

In cryptography, searchable encryption is a type of encryption that allows for the search and retrieval of data within encrypted data without the need to decrypt the entire data set. Using encryption/searchable encryption technologies, cloud secure architecture [13, 14] enables data retrieval in a secure manner while also allowing for the search process to be carried out in the form of encrypted data.

2.3 Homomorphic Encryption

When decrypted, homomorphic encryption [15, 16] is a type of encryption technique, in which a computation is performed on the cipher text and the result thus generated matches the result of the operation performed on the plaintext when the cipher text is decrypted. In general, the purpose of the homomorphic technique is to maintain the integrity of data while it is being transmitted over the cloud.

2.4 Encryption Based on Attribute

Attribute-based encryption (ABE) [17] is a one-to-many encryption technique that uses public-key cryptography to allow users to encrypt and decrypt data based on the attributes of the users. ABE schemes are classified into two categories: key-policy ABE and cipher text-policy ABE [18, 19, 20].

2.5 Hybrid Encryption

Hybrid encryption [21, 22] is a type of encryption that combines the strengths of two or more encryption systems in order to benefit from the advantages of each type of encryption separately.

3 Proposed Methodology

Security and privacy of information are becoming increasingly important concerns for everyone as more and more businesses and individuals choose to store their data in the cloud. There should be no doubt about it: Data files are primarily user-centric when it comes to encryption and decryption. This means that only legitimate users are permitted to upload or download files, and that users have the ability to specify

whether a file can be shared with other users. Data security in a cloud environment is comprised of two components that must be taken into consideration. In the beginning, data security may be a concern when data is transferred into the network after being collected from the user site through any web-based application, especially when the data is sensitive. A security concern could arise when data has already passed through the network and is about to be stored on a cloud-based storage device. The primary motivation for the proposed work is to address the second concern, which is security concerns about the data file at the cloud end, while the data is being stored in the cloud disc. The proposed work will address both concerns. In order to maintain cloud storage security, researchers have provided the following skeleton of the proposed work, which is hybrid in nature and consists of three stages.

The Caesar Ciphering technique [22], for example, is used in the first stage, as depicted in Fig. 1, and it provides an initial level of security while also increasing the efficiency of the system, without question. Second, and more specifically, the proposed work is concerned with a newly designed encryption algorithm that is based on the symmetric cryptographic concept, which is described in greater detail further down this page (block based). Because of the use of a 128-bit block size for encryption purposes, a higher level of security can be achieved with this work than with other works. This 128-bit block size is encrypted at the same time with the help of an encrypted key that has a 128-bit size as well, which is also encrypted. An encrypted key for use in the cloud is generated by performing an XOR operation on the private key of the user and the secret key of the cloud. Thus, the newly designed encryption algorithm provides a double layer of data security, which is a significant improvement over the previous algorithm. To summarize, it should go without saying that when it comes to data security, cryptographic encryption techniques play a critical role in protecting sensitive information. The user who attempts to access these cloud-based storage data should, however, be verified to ensure that they have the appropriate authentication rights. It is critical to have cloud data security measures in place because they require authentication or verification of the user before granting access to the cloud data. Because of this, the third stage of the newly proposed work focuses on the authentication of cloud users through the use of Attribute-Based Cryptography (ABC) techniques in the cloud, as opposed to the previous stage. With the help of this technique, the algorithm generates an attribute associated with the cipher text, and the authentication of the user's requisite is handled in accordance with the attribute generated by the algorithm. As long as the user meets this requirement, the newly proposed cloud-based security system will also perform a search for the aforementioned key.

The proposed methodology assesses five major objectives, confidentiality, authentication, integrity, access control, and non-repudiation. We are focusing on all aspects of security in order to improve the overall performance of cloud computing at all stages of development. The proposed solution is primarily divided into three major objectives, which are briefly summarized as follows: Privacy and confidentiality of personal information.

The proposed solution points out that hybrid encryption algorithms have the potential to both increase the level of security while also improving the level of privacy.
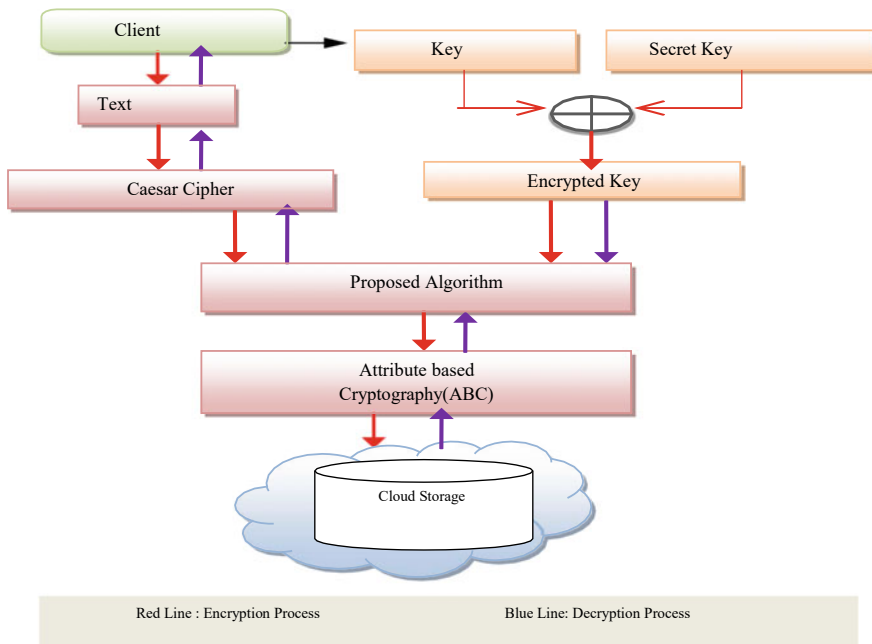
Fig. 1 Block diagram of proposed concept

As symmetric cryptography, this paper proposed the use of the RC6 and Blowfish algorithms, while as asymmetric cryptography, ECC and RSA were proposed.

3.1 Authentication, Authorization, and Access Control

The proposed authentication solution would assist users in identifying themselves in a more accurate and secure authentication manner. With the help of this field, you can improve the level of user authentication and authorization by experimenting with different approaches.

This model's goal is to improve the performance of authentication by proposing certain techniques that are followed, such as the following:

i. The username and password technique
ii. The verification of security questions approach
iii. The use of a security token for verification
iv. One-time password verification for mobile devices
v. IP and MAC address verification
vi. Authentication systems based on the Kerberos protocol.

This work investigates the fact that there are a variety of scenarios in which a user will attempt to log into a system. In this section, a few of the most common cases are considered in order to propose an authentication layer in a specific situation.

3.2   Access Control

In this paper, a hybrid approach for healthcare cloud computing is used, which includes RBAC and ABAC access control, as well as encryption techniques such as ECC and Blowfish [23].

The following are some examples of how health-related information can be threaded together:

- Intruders who pose a human threat can be identified as hackers.
- Natural disasters such as earthquakes, fires, and other calamities pose a threat.

  Problems and failures related to technology, such as system damage and crashing.

  By analyzing the entire work, its primary goal is demonstrated through the use of some mitigating approaches that are used to diagnose the existing work.

  i.   To simulate the mapping of role- and attribute-based access control policies and procedures.
 ii.   To diagnose the problem of information misdirection caused by insecure and sensitive information.
iii.   It determines which roles can be accessed by which attributes and which attributes cannot.

The entire work is donated for the purpose of maintaining privacy in the information.

The use of cloud computing in the healthcare sector has the potential to fundamentally alter the way that medical treatment and research are conducted and conducted. However, it has also presented a slew of difficulties and restrictions for the same. The privacy and security of the electronic medical record (EMR) are the most significant challenges associated with the implementation of cloud computing in health care. In the healthcare sector, the security frameworks that are currently in place are not very efficient or secure in the face of the ever-increasing threat of cybercrime. The compromised medical record may be severely exploited by malicious stakeholders, resulting in a medical blunder that could be felt all over the world if it is made public. As a result, the researchers are extremely concerned about the security of the massive amount of medical data they have collected A hybrid technique will be implemented, and then permission will be granted to the user on the basis of a predetermined limit of threshold value will be established. On the basis of hybrid technique, which is a combination of Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC), these threshold values are set, which will determine whether the user is an attacker or a genuine user on the network.

At the end of the decryption process, the authenticate user will be informed that even chunks will be decrypted using the ECC algorithm and odd chunks will be decrypted using the Blowfish algorithm, according to the information provided. The following is the procedure to be followed in order to implement the hybrid security framework for EMR:

- Over an insecure Internet connection, the user uploads original data (text and image) and then verifies the relative fake data cloud that has been created.
- Data is divided into chunks in binary form, which is called bifurcation.

Then, using the ECC and Blowfish techniques, the original and fictitious data will be encrypted using the same method.

In addition, these cryptographic techniques will be applied at random to any random data that is generated. It will be used to shuffle data in preparation for further processing. Finally, the entire study addresses a security model for a specific security later on, as well as the algorithms that are recommended. The following are all of the recommended algorithms and layers given in Table 1.

4 Experimental Analysis

Here, we can see how well the hybrid algorithm proposed performs in terms of cipher text size, encryption time, and decryption time taken by the hybrid algorithm. Using different sizes of paper, we can determine how long it will take to complete a calculation? The hybrid algorithm that has been proposed is also compared to the algorithms that are currently in use. Different cryptography techniques such as RC6, AES, BLOWFISH, ECC, and RSA are used to compare and evaluate the performance of encryption time and decryption time in milliseconds. Experimental analysis is carried out on data of various sizes including 1, 10, 100, 1000, and 10,000 KB, given in Tables 2 and 3.

Experimental analysis enhances working on different size of data like 1, 10, 100, 1000 and 10,000 KB. A comparison of different execution time is given in Table 4, and Fig. 2 shows the comparison of different cryptographic algorithm.

Table 2 Encryption time (ms): comparison of encryption time

Data (KB)	AES	RC6	Blowfish	RSA	ECC
1	3.1	2.4	27	111	86
10	18	17	41	213	116
100	94	62	54	514	391
1000	214	164	65	2354	857
10,000	1855	1324	103	4570	1669

Table 3 Decryption time (ms): comparison of decryption time

Data (KB)	AES	RC6	Blowfish	RSA	ECC
1	4	4	28	187	108
10	22	18	47	287	155
100	102	77	64	797	455
1000	246	196	76	2657	957
10,000	1987	1676	124	4970	1870

Table 4 Execution time (in seconds): comparison of different execution time

Data (KB)	Traditional system	Proposed model
1	3	4.5
10	19	22.56
100	95	96.25
1000	285	312
10,000	2252	2312

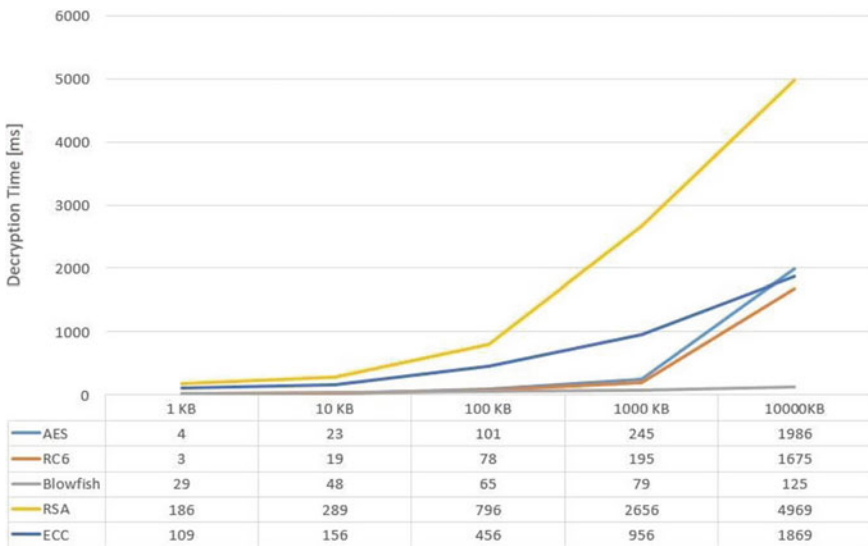|	1 KB	10 KB	100 KB	1000 KB	10000KB
AES	4	23	101	245	1986
RC6	3	19	78	195	1675
Blowfish	29	48	65	79	125
RSA	186	289	796	2656	4969
ECC	109	156	456	956	1869

Fig. 2 Comparison of different cryptographic algorithm

5 Results

The complete work concludes that security plays significant role to improve the confidentiality of system along with overhead for memory and computation also. Because of the use of cryptographic techniques, the purpose of this review was to

draw attention to the confidentiality difficulties associated with cloud computing. It is also concluded that by removing mobile verification and Kerberos from the private network, we can save a substantial amount of time and effort. Consequently, the integration of IP and MAC address verification aids in maintaining a higher level of security while simultaneously reducing the overhead of cost and time by eliminating the need for mobile one-time password (OTP) authentication. The 3A's (Authentication, Authorization, and Access Control) are being used in this approach to implement a multidimensional approach and overcome the issues that have been identified in the existing work regarding data security and privacy protection on a two-dimensional scale when using cloud computing.

6   Conclusion

This paper mainly focuses on solution to only increase level of security to insure proof of identity and access rights but also help to differentiate sensitive data and general purpose data. It increases little computation time and enhanced execution time is negligible against the level of security which has been increased due to AAA model. This solution can be implemented for any public platform, who having sensitive data. This model can also be being integrated with e-commerce and other platform to increase the level of security.

References

1. Liu F, Tong J, Mao J, Bohn R, Messina J, Badger L, Leaf D (2011) NIST cloud computing reference architecture. US Department of Commerce, Gaithersburg, MD
2. Mell P, Grance T (2011) The NIST definition of cloud computing. Special publication 800-145. US Department of Commerce, Gaithersburg, MD
3. Rimal BP, Choi E, Lumb I (2009) A taxonomy and survey of cloud computing system. In: International joint conference on INC, IMS and IDC. IEEE
4. Tsai W-T, Sun X, Balasooriya J (2010) Service oriented cloud computing architecture. In: Seventh international conference on information technology. IEEE
5. Sajid M, Raza Z (2013) Cloud computing: issues and challenges. In: International conference on cloud, big data and trust
6. Ramgovind S, Eloff MM, Smith E (2010) The management of security in cloud computing. In: Proceedings of information security for South Africa (ISSA). IEEE, 128–142
7. Hashizume K, Rosado DG, Fernández-Medina E, Fernandez EB (2013) An analysis of security issues for cloud computing. J Internet Serv Appl (Springer)
8. Dahal S (John W. Ritting house and James F. Ransome) (2010) Security architecture for cloud computing platform. CRC Press, Taylor & Francis Group, Boca Raton, FL. ISBN: 978-1-4398-0680-7
9. Eletriby S, Mohamed EM (2012) Modern encryption techniques for cloud computing randomness and performance testing. In: ICCIT 2012. Al-Sabri HM, Al-Saleem SM (2013) Building a cloud storage encryption (CSE) architecture for enhancing cloud security. IJCSI Int J Comput Sci 10(2)

10. Patidar G, Agrawal N, Tarmakar S (2013) A block based encryption model to improve avalanche effect for data security. Int J Sci Res Publ 3(1)
11. Tiwari R, Sharma M, Mehta KK, Awasthy M (2020) Dynamic load distribution to improve speedup of multi-core system using MPI with virtualization. Int J Adv Sci Technol 29(12s):931–940. ISSN: 2005–4238
12. Tiwari R, Sharma M, Mehta KK (2020) IoT based parallel framework for measurement of heat distribution in metallic sheets. Solid State Technol 63(06):7294–7302. ISSN: 0038-111X
13. Lin M-PP, Hong W-C, Chen C-H, Cheng C-M (2013) Design and implementation of multi-users secure indices for encrypted cloud storage. In: international conference on privacy, security and trust. IEEE (Trend Micro, Taiwan)
14. Gambhir S, Rawat A, Sushil R (2013) Cloud auditing: privacy preserving using fully homomorphic encryption in TPA. Int J Comput Appl 80(14)
15. Wang C, Chow S-M, Wang Q, Ren K, Lou W (2013) Privacy-preserving public auditing for secure cloud storage. IEEE Trans Cloud Comput 62(2)
16. Zhu S, Yang X, Wu X (2013) Secure Cloud file system with attribute based encryption. In: IEEE international conference on intelligent networking and collaborative systems
17. Yang C, Lin W, Liu M (2013) A novel triple encryption scheme for hadoop-based cloud data security. In: IEEE international conference on emerging intelligent data and web technologies
18. Tiwari SK, Rajput DS, Sharma S, Neogi SG, Mishra A (2020) Cloud virtual image security for medical data processing. Math Model Soft Comput Epidemiol 317–345
19. Tiwari SK, Neogi SG, Mishra A (2021) Design and implementation of secure system for virtual machine image in cloud computing. Des Eng 15044–15054. http://www.thedesignengineering. com/index.php/DE/article/view/4787

Multi-domain Opinion Mining: Authenticity of Data Using Sentiment Mining

Bipin Kumar Rai, Satyam Gupta, Shubham Dhawan, and Nagendra Nath Dubey

Abstract With all the advancement of Internet telecommunications, a huge quantity of data is usually present for web users. Users make use of the resources offered and provide comments, resulting in the generation of more useful data. Because of the large number of people's perspectives, beliefs, comments, and recommendations obtainable via online resources, it's critical to study, examine, and categorize their viewpoints in order to make better decisions (ChandraKala and Sindhu in Opinion mining and sentiment classification: a survey, 2012). E-commerce websites have a significant impact on our day-to-day life. Whether it's a handset or a vehicle, the average customer depends heavily on public feedback and comments provided (by others) on the Internet to learn about any goods prior to making a choice. The goal of our research study is to robotize the number of digital end-user ratings for each item or brand and assess those available evaluations for attitudes expressed regarding certain aspects. This process involves filtering irrelevant and unhelpful thoughts from multiple sources for checking their reliability and quantifying the feelings of thousands of user reviews.

Keywords Sentiment analysis · Sentiment mining · Data analysis · Data extraction · Web scraping · Data authentication

1 Introduction

Ideas, views, moods, recommendations, and good/bad are all examples of sentiments. By evaluating a huge number of articles, the Data Exploration job tries to collect the writer's thoughts reflected in favorable or negative remarks, queries, and requests. Opinion inspection is a type of human behavior research that involves extracting

B. K. Rai (✉) · S. Gupta · S. Dhawan · N. N. Dubey
Department of Information Technology, ABES Institute of Technology, Ghaziabad, Uttar Pradesh 201009, India
e-mail: bipinkrai@gmail.com

N. N. Dubey
e-mail: nagendra.dubey@abesit.edu.in

user sentiment and emotion from plain text. Opinion mining is another name for sentiment analysis [2].

Numerous research on the impact of Internet suggestions on user decisions has previously been conducted [3]. But, from the point of view, getting the point of such disparate sets of data (articles, journals, and customer posts) scattered over disparate, unconnected platforms is a difficult challenge. As an outcome, an automated system is required to collect, process, summarize, and illustrate such opinions to assist manufacturers and enterprises in enhancing their goods based on user feedback [1].

Despite the fact that there has been a lot of effort done on identifying themes, these lines of research have mainly concentrated on locating and evaluating document difficulties. There is no sentiment analysis within content that restricts the utility of the extraction outputs [4]. Earlier research looked at the challenge of sentiment identification at differing stages (from keywords to paragraphs to articles). None of them, however, can model a combination of themes and sentiment categorization, leaving the findings less useful to consumers [4]. Our idea is to collect feedback from different sources so that when we form an opinion about some product (like an article, movie, music, etc.), we are not biased by any one type of news/info. So, to do that, we are trying to make a web app that can analyze the opinions of a vast number of users of the product from different sources and give it to one place.

2   Related Work

Kumar et al. [3], where they evaluated the literature in reputable business management publications that employed text mining approaches, including Opinion Analysis, Topic Modeling, and Natural Language Processing (NLP). Further, to explore the prevalent themes and linkages, they used visualization techniques for concept identification and data mining.

Vashishtha and Susan [5], Using the Sent WordNet lexicon along with some techniques of fuzzy linguistic hedges, introduced an unsupervised sentiment classification model that thoroughly formulates phrases, computes their senti-scores and polarities. Ridhwan and Hargreaves [6], here they used sentiment classification and topic modeling on tweets concerning COVID-19 effects in Singapore from February 1, 2020, to August 31, 2020, for this research. They did this by utilizing the Python module 'SNSCRAPE' to gather tweets about COVID-19 and geolocate it as Singapore.

Zarindast et al. [7], here to provide a research plan, in the smart lighting industry, they look at current literature and user perceptions. For this study, they use two data sources: The first one is the lighting research literature from the phase of research, and the second one is Amazon reviews of two novel lighting products from the phase of development, indicating customers' impressions.

Neogi et al. [8], here they utilized tweets and data from Twitter, a microblogging platform to gather information on farmers' protests to comprehend public mood on a global scale better. Depending on even a tally of around 20,000 protest-related tweets, they deployed algorithms to identify and evaluate the feelings.

Obembe et al. [9], during the initial phases of COVID-19, this research article explores public tourist responses to crisis messaging. The model of social-mediated crisis communication is used in this study to explore the crucial factors influencing general attitude during the early stages.

Birjali et al. [10], this article provides a comprehensive examination of sentiment classification analysis methodologies, problems, and trends to provide investigators with a worldwide overview of sentiment classification analysis and related topics.

Aljuaid et al. [11], various machine learning-based methods are examined in this study to assess the sentiments of in-text references. The results of the sentiment classification are then utilized to calculate positive, negative, and neutral reference scores.

Arun et al. [2], viewpoint analysis based on Tweets, which includes categorizing positive and negative opinions from Twitter posts as well as some deep extraction of favorable and unfavorable terms. We can easily determine the pros and cons of those tweets using this methodology.

Researchers here (i) study previously researched investigations on machine learning-based sentiment classification analysis to provide the foundation and (ii) review and analyze the existing and anticipated obstacles related to this research issue.

Tan et al. [12] describe the broad framework of the suggested Sentiment Monitoring system as well as a brief overview of its important components. Furthermore, the Sentiment Analyzer's sentiment composition criteria were described.

3   Proposed Solution

Multi-domain Opinion Mining is about the collection of data from different sites for the same product to create an unbiased opinion of the product. Because these sites contain massive amounts of data in the form of forums, comments, etc., they have, so to process this massive amount of data we are proposing this solution. In our solution, we are trying to create a web app that goes to different sites to browse through their comments and forum to make an opinion of the product as they cater to different audiences which contain different opinions as to the crowd changes with the site. So to create an unbiased opinion, we can take the opinion of different sites into consideration. The above process can be described through Figs. 1 and 2.
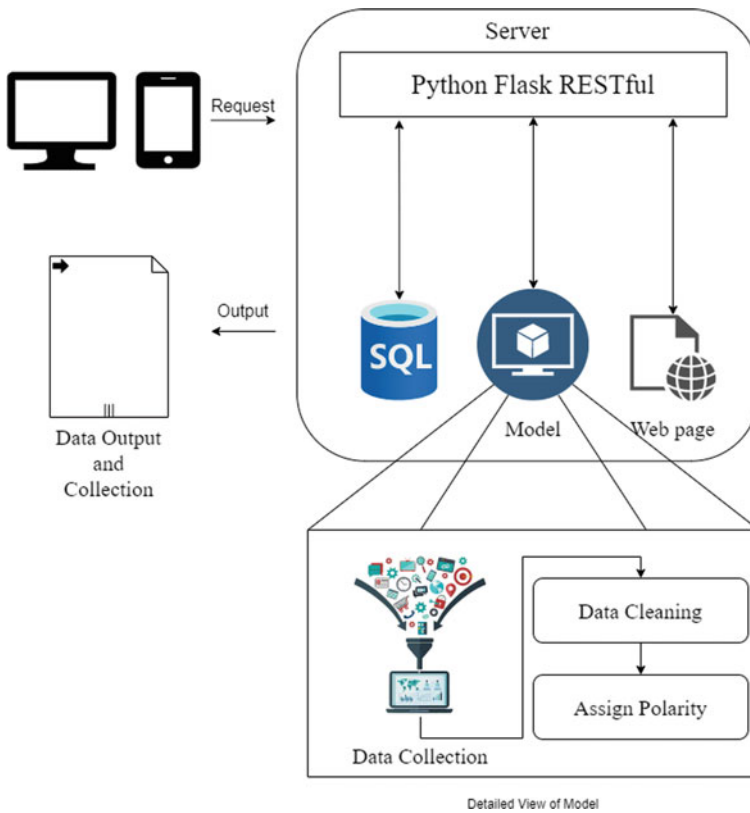
Fig. 1 Architecture: multi-domain opinion mining: authenticity of data using sentiment mining

3.1 Architecture

Server consist of four parts.

Python Flask RESTful Framework: It gives a framework to the server which connects all the components of the server like database, ML model and web pages. It also gives proper functionality to the site.

Web Page: These are the pages which collectively make the front end for the user who is visiting the side.

Model: It is the ML model which assigns to the data by using some steps like data collection, data cleaning, and assigning polarity to the processed data.

Back end: It is basically a database which stores the results and data collected by the model for future reference.

Fig. 2 Use case diagram: multi-domain opinion mining: authenticity of data using sentiment mining
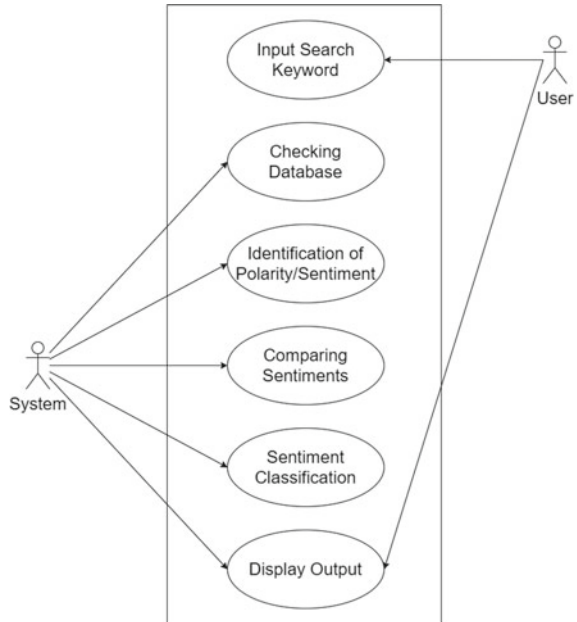
3.2 Use Case

Input Search Keyword—To access our web app, users first need to log in to our site and then enter their desired keywords related to the name of the product they require.

Checking Database—In this phase, the back end will check if the keyword or anything related to the keyword entered by the client is present in the database or not. If the keyword matches with the database, then further processing will be done, and if it is not present in the database, then APIs will be hit for data collection from different sites, and other processing will be done.

Identification of Polarity/Sentiment—In this built-in module, data is automatically preprocessed by filtering out unusual words and noises before further processing. Then only the filtered data remains. Now we can assign polarity to the data, and through the overall polarity, we can label the information whether the data is favorable or not.

Comparing Sentiments—After identifying sentiments, we compare product reviews from different sites and check the reliability and authenticity of data.

Sentiment Classification—With our Sentiment classifier, we can label or classify each keyword given by the client as positive, negative, and neutral sentiment in a few minutes.

Display Output—After comparing sentiments from different sites, we display the result for the clients, including the comparison of product reviews of various sites.

4 Methodology

Information Collection: The very first step gathers information from the social media sites such as LinkedIn, Facebook pages, Google, and blogs. This information is in an unstructured format. Thoroughly analyzing the text is difficult.

Text Preprocessing: Before analyzing the data, this stage is used to clean it up such that non-textual elements and useless crap are removed.

Sentiment Identification: This step has looked at the extracted phrase viewpoint. Subjective lines have greater sentiment, as they incorporate retrained beliefs, opinions, and reviews. Facts and factual information were removed that are considered as objective sentences.

Sentiment Stratification: It divides the statement into three polarities: (a) positive, (b) negative, and (c) neutral.

Sentiment Analysis Approaches: The basic stage in sentiment text categorization is to choose and recover text characteristics. The mining and selection of text characteristics is the initial stage in the Sentiment classification challenge. Major stress and recurrence, parts of speech (POS), and viewpoint idiomatic expressions are some of the components [9].

Recurrence and Duration of Appearance: Their distinguishing characteristics are specific words and their frequency counts.

Parts of Speech (POS): Finding an adjective is a critical view in tagging each word to its particular components of speech.

Terms and Phrases that Express an Opinion: Good or terrible, hate or like, and other terms are regularly employed to communicate opinions, while some representations offer viewpoints without using opinion words.

5 Results

Multi-domain Opinion Mining: Authenticity of Data Using Sentiment Mining introduces a new concept of the automated method that proposes to get the opinion analysis of data sets used to conduct a demonetization study of multiple sites, and expression categorization technique that is complete in its formulation, clean the data, remove the garbage comments, calculates their polarity and senti-scores.

Table 1 Rating of products on different e-commerce sites

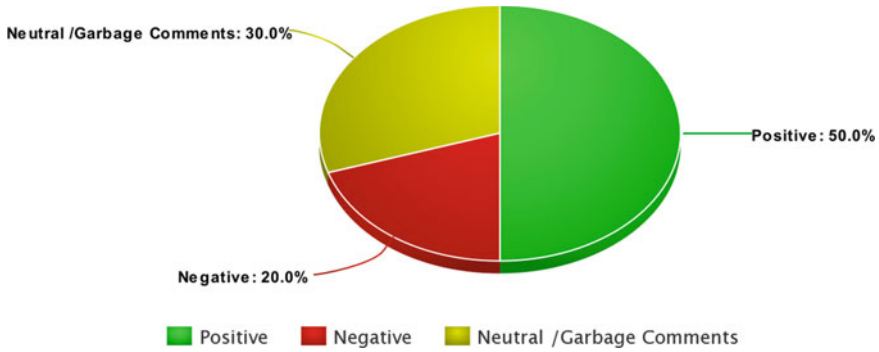| Products	Amazon rating	Flipkart rating	Others rating	Average
iPhone 13	4.6	4.7	4.5	4.6
MI	3.8	4.0	3.2	$3.7(\approx 3.666)$

Fig. 4 Pie chart of comments distribution

For example, we have taken two products' reviews from different sites as shown in Table 1.

Individuals and enterprises may use digital platforms and social networking sites to communicate ideas, opinions, and suggestions regarding services and goods. Many users and competitors have been known to abuse these systems by submitting phony reviews in order to harm a company's reputation. A pie chart of comment distribution of different attributes has been shown in Fig. 4. Such techniques may have a negative influence on a company's bottom line. Negative propaganda recognition and risk assessment utilizing text mining and natural language processing (NLP) might be researched in the future to gain a better knowledge of user demands in the service and communication fields.

To identify potential mismatches between research and development, we used data mining and aspect-based opinion classification. In terms of direction, date, and polarity of diverse sources of data, we evaluated and analyzed the themes stated in Table 2. Our Model Multi-source Opinion Mining takes data from almost all data sources of previous models.

6 Conclusion

To achieve benefits in diverse fields, the majority of the literature uses descriptive statistical analysis on review sites or social media posts. According to our comprehensive literature analysis, studies on the implementation of text mining have been widely

Table 2 Comparison with other models

Data taken from	Twitterizer	Multi-domain opinion mining	Flipkart sentiments analysis system	Snapdeal review management
Flipkart	✖	✓	✓	✖
Snapdeal	✖	✓	✖	✓
eBay	✖	✓	✖	✖
Twitter	✓	✖	✖	✖

published in a variety of journals. Several of these machine learning approaches have been covered in prior studies, such as Naive Bayes, Support Vector Machines, and Maximum Entropy. Discovering how consumers think can help researchers and developers improve modification percentages and give the answers they want.

In Table 1, the overall review/senti-score of iPhone 13 is 4.6 which is practically and theoretically correct with respect to error margin, i.e., authentic, but the overall review/senti-score of MI android mobiles (i.e., $3.7 \approx 3.666666$) varies gradually in different sites, i.e., not authentic or we can say that by analyzing the data that different e-commerce sites cater the different type of consumer crowds for a different type of need for the same product.

We intend to improve our methodologies in the future by incorporating semantic analysis. Furthermore, distributed data processing can boost the efficiency of our system while examining large amounts of data. We can further extend our project and help in medical fields in the treatment of patients in day-to-day life [13–15].

References

1. ChandraKala S, Sindhu C (2012) Opinion mining and sentiment classification: a survey
2. Arun K, Srinagesh A, Ramesh M (2017) Twitter sentiment analysis on demonetization tweets in India using R language. Int J Comput Eng Res Trends, 4(6):252–258
3. Kumar S, Kar AK, Ilavarasan PV (2021) Applications of text mining in services management: a systematic literature review
4. Agarwal B, Mittal N, Bansal P, Garg S (2015) Sentiment analysis using common-sense and context information. Comput Intell Neurosci 2015. https://doi.org/10.1155/2015/715730
5. Vashishtha S, Susan S (2021) Highlighting key phrases using senti-scoring and fuzzy entropy for unsupervised sentiment analysis
6. Ridhwan KM, Hargreaves CA (2021) Leveraging Twitter data to understand the public sentiment for the COVID-19 outbreak in Singapore
7. Zarindast A, Sharma A, Wood J (2021) Application of text mining in smart lighting literature—an analysis of existing literature and a research agenda
8. Neogi AS, Garg KA, Mishra RK, Dwivedi YK (2021) Sentiment analysis and classification of Indian farmers' protest using Twitter data. Emerging Markets Research Centre (EMaRC), Wales, UK
9. Obembe D, Kolade O, Obembe F, Owoseni A, Mafimisebi O (2021) Covid-19 and the tourism industry: an early-stage sentiment analysis of the impact of social media and stakeholder communication (Leicester)

10. Birjali M, Kasri M, Beni-Hssane A (2020) A comprehensive survey on sentiment analysis: approaches, challenges, and trends. (Author links open overlay panel)
11. Aljuaid H, Iftikhar R, Ahmad S, Asif M, Afzal MT (2020) Important citation identification using sentiment analysis of in-text citations
12. Tan LI, Phang WS, Chain KO (2016) Rule-based sentiment analysis for financial news. In: IEEE international conference, Kowloon, China, Jan 2016. https://doi.org/10.1109/SMC35812.2015
13. Rai BK (2022) Patient-controlled mechanism using pseudonymization technique for ensuring the security and privacy of electronic health records. Int J Reliable Qual E-Healthc (IJRQEH) 11(1):1–15. https://doi.org/10.4018/IJRQEH.297076
14. Rai BK (2022) Security challenges and solutions for healthcare in the internet of things. In healthcare systems and health informatics pp. 235–246. CRC Press
15. Rai BK (2022) Ephemeral pseudonym based de-identification system to reduce impact of inference attacks in healthcare information system. Health Serv Outcomes Res Method 1–19. https://doi.org/10.1007/s10742-021-00268-2

A Novel Low-Power NMOS Schmitt Trigger Circuit Using Voltage Bootstrapping and Transistor Stacking

S. Siva Kumar, Seelam Akhila, T. Ashok Kumar Reddy, A. Krishna Chaitanya, and G. Charan Kumar

Abstract This research paper investigates and experiments with a low-power Schmitt circuit based on transistor stacking. In the suggested VB-ST circuit, only an NMOS transistor is employed, which helps to reduce the ageing effect of the circuit, particularly Negative Bias Temperature Instability (NBTI), as well as the power consumption by employing the transistor stacking technique. In addition, the suggested Schmitt trigger is radiation toughened particle tolerant. With a supply voltage of 0.4 V, the suggested Schmitt trigger was constructed using a 45 nm Technology file and Tanner EDA tool.

Keywords NBTI · VB-ST · Power consumption

1 Introduction

As technology advances, CMOS devices become more susceptible to radiation-induced single event effects (SEE), which reduce the circuit's noise immunity. SEE occurs primarily as a result of reduced supply voltage and node capacitance, with single event upset (SEU) in storage elements being the most prevalent SEE. The SEU flips the logic states in digital devices, resulting in system failure. Due to the single event transient effect, combinational circuits may also create voltage glitches at the circuit's output (SET). The alpha particles and cosmic neutrons produced, especially in the terrestrial and space environments, trigger these SETs. However, when particles collide with a digital circuit node in the space environment, a logical error occurs, which is referred to as a soft error. For fault-tolerant circuits, the critical charge Qcrit quantifies the soft error rate at the most sensitive node in the circuit and should be as low as possible. The critical charge is the smallest quantity of collected charge at the circuit's sensitive node during a particle impact that is sufficient to affect

S. S. Kumar (✉) · S. Akhila · T. A. K. Reddy · A. K. Chaitanya · G. C. Kumar
Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, Kadapa, India
e-mail: sibyala.siva@gmail.com

the circuit's output state. As a result, it is evident that the critical charge should be as high as feasible in order to reduce the SER, and a lower SER enhances the circuit's soft error robustness.

For fault-tolerant circuit design, the inverter is the basic building block for most digital circuits, such as memory design and latches design. As a result, the inverter should be tolerant of BTI effects and soft error-resilient. The Schmitt trigger also provides superior stability in radiation settings, motivating us to create reliable inverter circuits based on ST. According to the supply voltage, a traditional CMOS inverter gives full swing from high to low. Due to the inclusion of PMOS and NMOS transistors in the CMOS circuit, it has a lower noise margin and is also influenced by the NBTI and PBTI phenomena. The PBTI, on the other hand, has received less attention, owing to its negligible impact on thin gate oxide. As a result, NBTI may have a greater impact on circuit performance than PBTI.

Only NMOS transistors are used in the voltage bootstrapped circuit, which not only eliminates the problem of aging, especially NBTI, but also reduces the noise margin and rail-to-rail voltage. Because NMOS transistors are present in the VB circuit, PBTI stress is also present, however PBTI may have a lower impact on NMOS devices than NBTI does on PMOS devices. As a result, for reliability study, we used BTI stress (NBTI + PBTI). For future VLSI systems, it offers the lowest leakage power and circuit delay, as well as decreasing the aging effect, making it ideal for reliable and low-power applications.

2 Literature Review

Soft error hardening enhancement analysis of NBTI tolerant Schmitt trigger circuit [1]. Effects of BTI and soft errors and techniques to improve them are studied in this the noise margin is improved. Design and Analysis of SEU Hardened Latch for Low power and high-speed applications [2]. Low power and noise margin are improved. Design for Ultra-Low-Voltage Operation [3] the idea of a sub-threshold voltage is investigated and for situations with stringent energy constraints, sub threshold digital circuit design has emerged as a low-energy alternative. Comparative Analysis of Schmitt Trigger with AVL (AVLG and AVLS) Technique Using Nanoscale CMOS Technology [4]. The CMOS device is used to improve speed, power dissipation, size, reliability, and hysteresis performance. The most effective method is to reduce power consumption and improve compatibility with low voltage power supply and analog components, as Schmitt trigger did. In this paper, the area of the circuit increases. Analysis and Design of a Low-Voltage Low-Power Double-Tail Comparator [5] in this the dynamic comparator shows that both the power consumption and delay time are reduced. High-speed low-power comparator for analog to digital converters [6]. Although the number of transistors is more, the latency and power consumption are lower. A low-power, high-speed two-stage dynamic comparator is presented in

this study. To lower the first stage's power consumption, the voltage swing of the comparator's first stage, the pre-amplifier stage, is limited to Vdd/2 in this circuit. This voltage swing limitation also provides a strong drive for the second stage during the evaluation phase, allowing the comparison speed to be increased and delay is more. A novel CMOS dynamic latch comparator for low power and high speed [6].

The comparators and suggested circuit were built and simulated in Tanner EDA suite utilizing 180 nm CMOS technology and a 1 V power supply voltage, and they show less power consumption and higher speed than traditional latched comparators.

3   Methodology

The power dissipation has become the critical issue in today's VLSI system. Because power dissipation is quadratically proportional to supply voltage (VDD), lowering VDD is the most effective technique to reduce it. However, as VDD decreases, circuit delay increases, lowering performance. At the same time, by lowering the threshold voltage (VTH), performance can be maintained, but sub-threshold leakage power grows exponentially. As a result, VDD and VTH must be tuned in order to provide the desired performance and low power consumption. To achieve low power consumption and to provide desired performance in this paper, we are implementing the circuit with Transistor Stacking.

The node capacitance and supply voltage are likewise aggressively scaled. As a result, the maximum number of charges that may be stored on a node is limited. As a result, logic architectures are more vulnerable to soft errors and external noise, such as alpha particles. When alpha particles impact a sensitive node in a digital design, the source/drain diffusion regions may form a secondary carrier, which causes glitches or transient faults (TF). The TF will appear as an electrical pulse when the amount of gathered charge is high, which is referred to as a single event transient (SET). The SET would spread across combinational logic and could cause problems in sequential circuits such as latches and flip-flops. If the deposited charge at the most sensitive node of the sequential circuit exceeds the critical charge at that sensitive node, the stored value has a chance to be flipped. A single event upset is the term for this phenomenon (SEU). The degrading effect Positive Bias Temperature Instability (PBTI) is beginning to play an essential role with the development of high-k gate oxide materials.

It has a major influence in this technique, especially when combined with the still-effective Negative Bias Temperature Instability (NBTI). Transistor stacking is a strategy for reducing leakage power.

When two or more series transistors are turned off, the leakage current reduces, which is known as the stack effect or self-reverse bias effect. The transistor stacking

Fig. 1 Stacked transistor
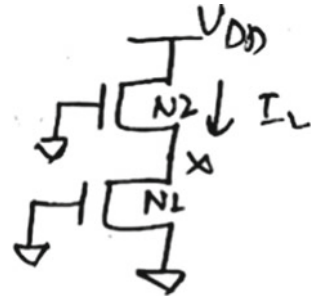
effect takes use of the subthreshold current dependence, and increasing the transistor's source voltage versus reduces the subthreshold leakage current exponentially. By increasing the number of transistors connected in a stack arrangement, more leakage power can be saved. Figure 1 shows how the transistor is stacked.

Process Flow

We are focusing on leakage power in our suggested solution in addition to constructing a BTI soft error tolerant circuit, which is critical in lower nanoscale technologies.

Figure 2 depicts the basic circuit we are implementing. This is the first step in the implementation process, and it involves the designing of the circuit which estimates the area based on the number of transistors we are using. Further, we have to take the inputs in the form of bits (1100). Further, the formulation of the Transient analysis and the power is carried in the process.

4   Results and Analysis

The proposed methodologies are implemented through the use of the TANNER EDA tool version 13.0 with 45 nm technology file which is computer- aided design program. The purpose of an EDA tool is to aid in the design and verification of a circuit's operation by numerically solving the circuit's differential equations. Circuit designers can use the simulation results to verify and fine-tune their designs before submitting them for fabrication. Figure 3 depicts the waveform of the Schmitt trigger with respect to the input.

The results are tabulated in Table 1. The power consumption in the proposed Schmitt trigger is reduced compare to existing Schmitt trigger circuit. The area of the proposed Schmitt trigger increased compared to the existing circuit because of the increase in the number of transistors. The circuit delay also decreases in the proposed Schmitt trigger.
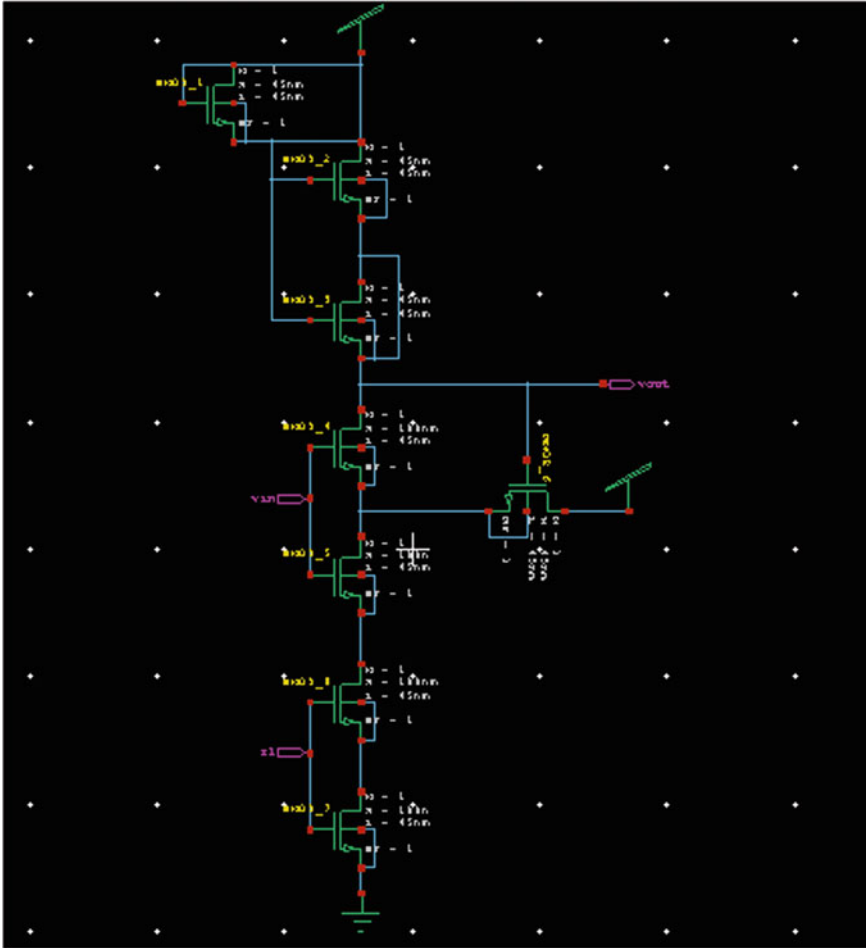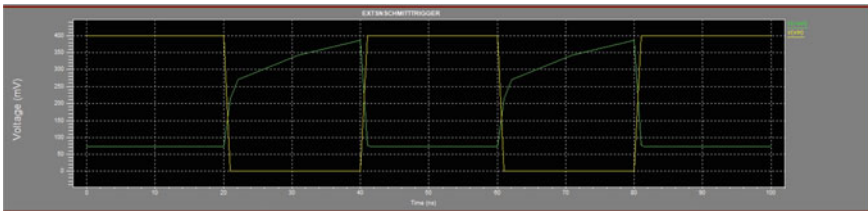
Fig. 2 Proposed circuit

Fig. 3 Waveform of proposed Schmitt trigger

Table 1 Power consumption and delay time of Schmitt trigger

Name	Power (uw)	area	Delay (ns)
Existing Schmitt trigger	0.8	6	0.68
Proposed Schmitt trigger	0.5	8	0.65

5 Conclusion

A low-power, soft-error-tolerant Schmitt trigger is also designed in this study. The suggested circuit uses only NMOS transistors to lessen the NBTI effect and includes a feedback mechanism to improve robustness against external noise. The suggested VB-ST circuit also has the lowest leakage power and circuit latency, as well as a lower ageing impact, making it ideal for high-reliability, low-power applications.

References

1. Shah AP, Yadav N, Beohar A, Vishvakarma SK (2018) On-chip adaptive body bias for reducing the impact of NBTI on 6T SRAM cells. IEEE Trans Semicond Manuf 31(2):242–249
2. Wang Y, Enachescu M, Cotofana SD, Fang L (2012) Variation tolerant on-chip degradation sensors for dynamic reliability management systems. Microelectron Reliab 52(9):1787–1791
3. Schroder DK (2007) Negative bias temperature instability: what do we understand? Microelectron Reliab 47(6):841–852
4. Grasser T (2012) Stochastic charge trapping in oxides: from random telegraph noise to bias temperature instabilities. Microelectron Reliab 52:39–70
5. Rzepa G, Franco J, O'Sullivan B, Subirats A, Simicic M, Hellings G, Weckx P, Jech M, Knobloch T, Waltl M et al (2018) Comphy—a compact-physics framework for unified modeling of BTI. Microelectron Reliab 85:49–65
6. Bagatin M, Gerardin S, Paccagnella A, Faccio F (2010) Impact of NBTI aging on the single-event upset of SRAM cells. IEEE Trans Nucl Sci 57(6):3245–3250
7. Shah AP, Yadav N, Beohar A, Vishvakarma S (2019) SUBHDIP: process variations tolerant subthreshold Darlington pair based NBTI sensor circuit. IET Comput Digital Tech 13(3):243–249
8. Schrimpf R, Warren K, Weller R, Reed R, Massengill L, Alles M, Fleetwood D, Zhou X, Tsetseris L, Pantelides S (2008) Reliability and radiation effects in IC technologies. In: IEEE international reliability physics symposium, pp 97–106
9. Schroder DK, Babcock JA (2003) Negative bias temperature instability: road to cross in deep submicron silicon semiconductor manufacturing. J Appl Phys 94(1):1–18
10. Shah AP, Yadav N, Beohar A, Vishvakarma SK (2018) An efficient NBTI sensor and compensation circuit for stable and reliable SRAM cells. Microelectron Reliab 87:15–23
11. Shah AP, Rossi D, Sharma V, Vishvakarma SK, Waltl M (2020) Soft error hardening enhancement analysis of NBTI tolerant Schmitt trigger circuit. Microelectron Reliab 107:113617
12. Satheesh Kumar S (2019) Design and analysis of SEU hardened latch for low power and high speed applications. J Low Power Electron Appl
13. Calhoun BH, Chandrakasan AP (2007) A 256-kb 65-nm sub-threshold SRAM design for ultra-low-voltage operation. IEEE J Solid-State Circ
14. Saxena A, Akashe S (2013) Comparative analysis of Schmitt trigger with AVL (AVLG and AVLS) technique using nanoscale CMOS technology. In: Third international conference on advanced computing and communication

15. Babayan-Mashhadi S, Lotfi R (2014) Analysis and design of a low-voltage low-power double-tail comparator. IEEE Trans Very Large Scale Integr (VLSI) Syst 22(2):343–352
16. Khorami A, Sharifkhani M (2016) High-speed low-power comparator for analog to digital converters. AEU—Int J Electron Commun
17. Singh S (2015) A novel CMOS dynamic latch comparator for low power and high speed. Int J Microelectron Eng (IJME) 1(1)
18. Fang J, Sapatnekar SS (2013) The impact of BTI variations on timing in digital logic circuits. IEEE Trans Device Mater Reliab 13(1):277–286

Dynamic Channel Allocation in Wireless Personal Area Networks for Industrial IoT Applications

Manu Elappila, Shamanth Nagaraju, K. S. Vivek, and Ajith Gopinath

Abstract Industrial wireless networks gain a substantial growth in size in the global market. In the congested scenarios of the industrial IoT application instances of wireless personal area networks, it should have a medium access strategy that is efficient and works autonomously to provide a reliable channel by reducing packet collisions. Medium access protocols must consider properties of the links between devices before a node is allowed to access the shared medium. Characteristic metrics of the channel like link quality indicator, received signal strength indicator, and path loss distance have to be considered in the contention resolution process between the nodes. A fuzzy-based channel allocation algorithm is proposed with dynamic adaptation of contention window in channel access strategy of the MAC layer standard. As per the simulation results, the algorithm proposed showed better results in terms of network throughput and packet delivery rate.

Keywords IIoT · WPAN · Channel allocation · Fuzzy

M. Elappila (✉) · S. Nagaraju
Department of Computer Science and Engineering, CHRIST (Deemed to be University),
Bangalore, Karnataka 560074, India
e-mail: manu.elappila@christuniversity.in

S. Nagaraju
e-mail: shamanth.n@christuniversity.in

K. S. Vivek · A. Gopinath
Department of Mechanical Engineering, CHRIST (Deemed to be University), Bangalore,
Karnataka 560074, India
e-mail: vivek.ks@christuniversity.in

A. Gopinath
e-mail: ajith.gopinath@christuniversity.in

1 Introduction

Industry 4.0, the fourth industrial revolution (4IR), will transform most industries across the globe in the upcoming years with the technological developments of the recent past and rapid globalization. The first industrial revolution came with the invention of steam engines that succeeded human labor and the second industrial revolution switched operations using electric energy. The next industrial revolution was introduced with automation having aid from computers and the Internet [1]. In the near future, the superintelligence revolution based on the Internet of things (IoT), cyber-physical systems, and artificial intelligence (AI) will greatly change human intellectual labor.

Although many industries have started to implement 4IR, implementation of these in a full-fledged manner has always been a challenge. Recent studies and reports have shown how the productivity of companies can be boosted with this revolution but will have an impact on each associated process [2]. The current study anticipates the changes in the nature of work and the way employees in the industry think with this revolution. In addition to it, with machines that replace a major part of the labor force, there will be complete reconstruction in the labor market; new jobs will replace the existing ones. The movement toward the usage of robots is already affecting the low-skilled workers, and their jobs are under threat. The future market for jobs will be based on the nature of the work and whether those can be executed by skilled robots or not. Hence, technological developments despite having a positive future can lead to a lot of unemployment in the field but can open new doors to job opportunities with different skill-sets.

The associated technologies which will be driving Industry 4.0 are diversified where approaches like artificial intelligence is trying to revolutionize the manufacturing industry with advanced networks and data management. In a short period of time, operational technology or cyber-physical system devices will monitor, coordinate, and integrate information in real time [3]. AI, machine learning, cloud computing, and IoT together have the capability to make interactions of machines with humans at a better level. The digital platforms can simulate human behavior with the sufficient data available. The purchasing patterns of the crowd can also be predicted from the evolving models with high accuracy levels.

Internet of things has become an essential element for the current industrial networks. The data that is received from different objects and their surrounding is being analyzed in order to provide support for different applications [4]. IoT employs network devices with low power and restricted computing and communication capabilities to collect data. Most of the IoT networks are in the form of mesh topologies that are extremely effective with hyper-competitive topology structure and node characteristics [5]. According to Grand View Research, Inc., the size of global industrial wireless network will gain a substantial growth in 2025 and the increasing investments and technological proliferation in the IoT applications will drive the global market for next seven years [6, 7]. Industrial wireless networks are dense networks with sensors, actuators, and other communicating devices that may

get added over time during the operation phase. Hence, to minimize the network performance degradation in these ever-growing mesh networks, distributed logics with fuzzy-based MAC protocol can be applied that can autonomously improve the throughput by prioritizing the access to the medium for nodes with better channel characteristics.

The smart devices in the IIoT applications may generate data packets in repeated time intervals. Sensors in coal mine environment monitoring, production flow management, gas pipeline monitoring, etc., are examples of this type of data generation. Moreover, a set of sensor nodes are getting activated, gather the data, and then go to idle state. After that, another set of nodes are activated and do the same which manages overlaps in the target environment. Hence, many devices in a particular region communicate at the same time, and they may send the assembled data to the next hop toward the base station that connects to the outside Internet. Figure 1 shows such a sample topology. In typical machine-to-machine (M2M) communication instances and IIoT applications, there could be simultaneous transmissions happening. The nodes and end devices are located near to each other as they are either mounted on a machine or worn by human beings. In such applications, there are chances that the transmission range and carrier sensing range of those communicating devices to overlap each other as they are in the operating space of a machine or human. There may be numerous sender–receiver pairs in the transmission range of a single node that can create collisions at the MAC layer. In the congested scenarios of the IoT application instances of low-rate wireless personal area networks (LR-WPAN), it should have an efficient medium access strategy and hence an intelligent channel allocation. The existing methods use application data rate and parameters such as position and response time for contention window optimization. But, with these fast-changing and ever-growing networks, solutions that work autonomously and independent of the characteristics of the node have to be developed. And the channel-condition aware upper layer protocol as in [8, 9] has to be used for better performance. MAC protocols should consider the properties of links between the devices before allowing a node to access the shared medium. Characteristic metrics of the channel like link quality indicator (LQI) and received signal strength indicator (RSSI) have to be considered in the contention resolution process between the nodes. Hence, in this paper, it is tried to develop a dynamic channel allocation strategy among the contending nodes in an industrial IoT network scenario according to the channel characteristics at the MAC layer of the links between the nodes.

The remainder of this paper is organized as follows: Sect. 2 provides a brief overview of the related works. Section 3 presents the dynamic channel allocation mechanism with contention window adaptation. Section 4 illustrates the simulation and results analysis, and Sect. 5 addresses the conclusion of the work.
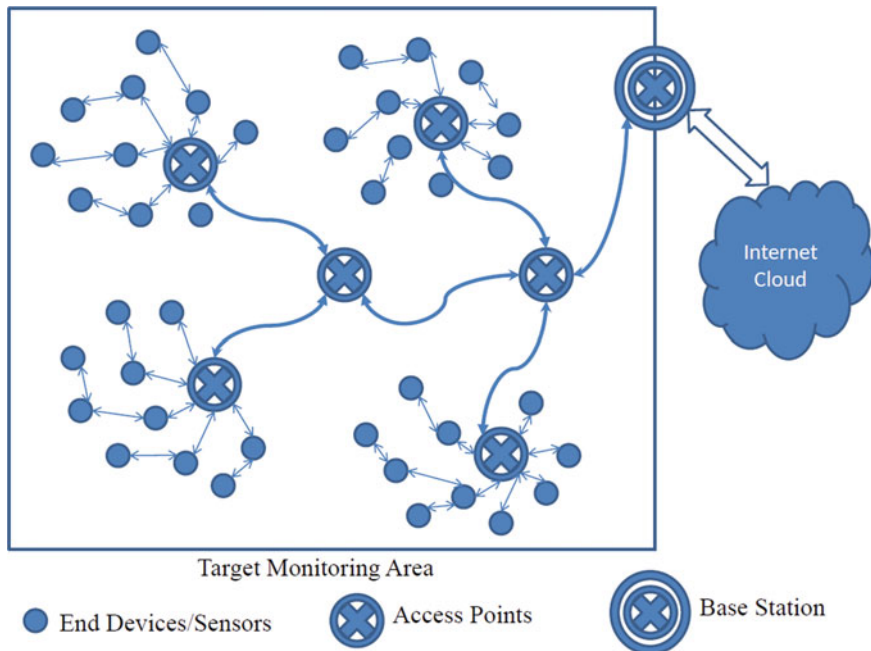
Fig. 1 A sample topology with many sender–receiver node pairs in an overlapped transmission range. Access points are deployed for coordinating the data transfer to the base station

2 Related Works

The medium access control sublayer and the physical layer characteristics of LR-WPAN are defined in IEEE 802.15.4 standard [10]. It uses both contention access periods (CAP) and contention-free periods (CFP) with guaranteed time slots as the MAC strategy. CSMA/CA is used in CAP to resolve contention between the nodes. Some previous works in the literature improve the contention process of LR-WPAN MAC protocols. DAH-MAC is proposed in [11] for single hop IoT-enabled MANETs. They use hybrid superframe structure to accommodate both voice traffic and data traffic. It uses contention window size adjustment in the CAP of the superframe for best-effort data traffic. For delay-sensitive voice traffic, adaptive time slot allocating TDMA-based CFP is used. In [12], fuzzy logic-based cross-layer MAC strategy is proposed. The back-off counter is calculated based on the number of previous back-offs and the application data rate. Authors in [13] try to propose a MAC improvement with analysis based on load, position, and response time. Contention window optimization is applied based on the cooperative load analysis. Back-off algorithm is the predominant component of all contention-based MAC strategies [14]. There are contention window adjustment techniques which use half feedback mechanism with

a shortcoming of CW diverging problem. Full-feedback back-off algorithm with dynamic CW adjustment is presented in [15]. Authors of [16] present an analytical model for evaluating the performance of IEEE 802.15.4 CSMA/CA. They evaluate the protocol with different sizes of the contention window and back-off exponent.

In WPAN protocol suit, IEEE 802.15.4 standard defines the lower layers PHY and MAC. The MAC sublayer of a PAN coordinator is responsible for generating network beacons. For all the nodes, their access to the physical radio channel is handled by the MAC sublayer. Other functionalities of the devices in a PAN like synchronizing with the beacon packets, association, and disassociation are governed by this layer. There are some other responsibilities also for MAC sublayer like handling the carrier sensing for channel access, providing and maintaining GTS mechanism, bearing the reliability of the link between two peer MAC entities, and supporting device security.

WPAN is used to transfer data over relatively short distances. There is little or usually no infrastructure involved in the network connections of WPAN as opposed to WLAN. This feature necessitates the implementation of small, power-efficient, inexpensive solutions for protocols for the devices of WPAN. IEEE 802.15 working group and the task group four specify the standard for the PHY layer and MAC sublayer for WPAN [10]. It mainly focuses on the networking structure with low-power consumption, low data rate, and low cost that operates on a personal operating space (POS) of 10 m with a raw data rate of 250 kbps that is high enough to satisfy the application needs. Carrier sense multiple access with collision avoidance (CSMA/CA) is used for the access to channel. For the transfer reliability, it defines a fully acknowledged protocol. Usually, it uses allocated 16-bit short address, or it can use 64-bit extended address. Allotment of guaranteed time slots (GTSs) is also included in the superframe structure for the contention-free data transmission. Energy detection (ED), clear channel assessment (CCA), and link quality indication (LQI) are also the part of the channel sensing and quality assurance implementation [17]. A system that conforms to this standard can have two different types of devices full-function devices (FFD) or reduced-function devices (RFD). FFDs can work as a PAN coordinator or coordinator, but RFDs cannot [18]. RFDs are intended for simple applications like sensing and transferring data, and relaying, and consequently, minimal resources and memory capacity are needed to implement them. The superframe structure optionally binds the channel time of a PAN coordinator. The transmission of a particular control packet called beacons locates the starting and ending of a superframe. A superframe will have an active period and an inactive period. During the period of inactivity, the PAN coordinator can move to the sleep state which consumes less power also known as low-power mode [19].

The MAC sublayer provides MAC data service as well as MAC management service. MAC data service transmits and receives MAC protocol data units. Functionalities and features of MAC sublayer include beacon management for the synchronization of nodes, channel access control, GTS management, frame validation, managing the association and disassociation of nodes with coordinators, acknowledged frame delivery, etc. The superframe structure of the MAC layer contains an active period and an inactive period, during which the coordinators can enter into a power-saving mode. The time slotted active period has a contention access period and

an optional contention-free period. Guaranteed time slots are there in contention-free period, during which nodes can reserve the time slots for future transmission. In the contention access period, channel is accessed by the nodes using CSMA/CA.

3   Methodology of Dynamic Channel Allocation Algorithm

The standard channel allocation technique in wireless personal area networks uses fixed contention window (CW) method in CSMA/CA. CSMA/CA is the mechanism which is responsible for doing carrier sensing for all the devices those are trying to transmit the packet to their corresponding receivers. Before the start of transmitting their signals, nodes will have to make sure that the channel is clear, which is achieved by performing a clear channel assessment.

There are two variants of the algorithm, and those are slotted and unslotted. If the superframe structure is in place, slotted CSMA/CA is used. The back-off periods in the slotted variant of the algorithm are needed to be in alignment with the specific time slots of the superframe structure. Sixteen slots of equal length of time divide the active portion of the superframe in beacon-enabled networks. Non-beacon-enabled networks use unslotted CSMA/CA to access channel. There is no superframe structure in those networks and hence no necessity for the back-off slot alignment.

When channel is indicated as busy by the clear channel assessment (CCA), the device will have to wait for some time before the next attempt. This time interval is random and will be multiple of the unit back-off. The standard algorithm uses three variables, and they are BE, NB, and CW. BE stands for back-off exponent, NB for number of back-offs, and CW for contention window length. As described above, each instance the algorithm detects the channel as busy, it needs to back off for some time. The value of BE decides the allowed range of this random period. It would be a multiple of the unit back-off period. This integer multiplier is a number between 0 and 2BE−1. NB is a variable that tells the number of times a device went to back-off period and again tried to access the channel. It keeps tracking the number of back-offs and reties. Each time a device backs off due to the busy medium, NB is incremented once which was initially zero as defined at the start of the algorithm. If the channel access is not acquired even after a certain number of attempts, the algorithm will simply drop the packet and quit from further attempts. And the MAC layer reports this channel access failure to the higher levels. This maximum number of attempts is specified in the MAC information base by *macMaxCSMABackoffs* attribute.

The contention window (CW) variable can be utilized to find the value corresponding to the number of back-off periods the channel is clear before sending the frame. That is, if the value of CW is two then the device starts transmitting after the medium sensed idle toward the end of two consecutive back-off periods. The CW is used in beacon-enabled slotted CSMA-CA algorithm [20].

In the dense IoT scenario where several nodes contend for the medium, we are trying to assign the channel to that node having higher link quality metrics. Instead of static CW, it has been found out a dynamic value for the CW using the fuzzy inference system. When multiple source nodes are trying to use the same shared channel, there should be some mechanism to provide the medium access to the node which has the higher values for its link quality metrics. By using this algorithm, we are trying to allocate the channel to a node that has better LQI and RSSI metrics. Fuzzy logic has been applied to calculate dynamically, the value for the CW for each node. Nodes that are having better link quality metrics, LQI and RSSI and D_{PL}, get assigned a low CW and hence, get the channel access first. D_{PL} is the path loss distance which is a derived entity that can be calculated from the received signal power. If a node has weak LQI and RSSI values and high D_{PL} value, then it may get high CW and may defer from the transmission of the data.

Fuzzy logic defines the relationship between inputs and outputs just like human control logic [21]. Mamdani method inference engine is being implemented in the system with the COG-based defuzzification. The membership functions for the fuzzy inference system are shown in Fig. 2 and the rule base is in Table 1.

Fig. 2 Membership function of fuzzy system

(a) CSI mapped into [0, 255]

(b) LQI mapped into [0, 255]

(c) D_{PL} mapped into [0, 255]

(d) Output, Link Survivability Factor (z)
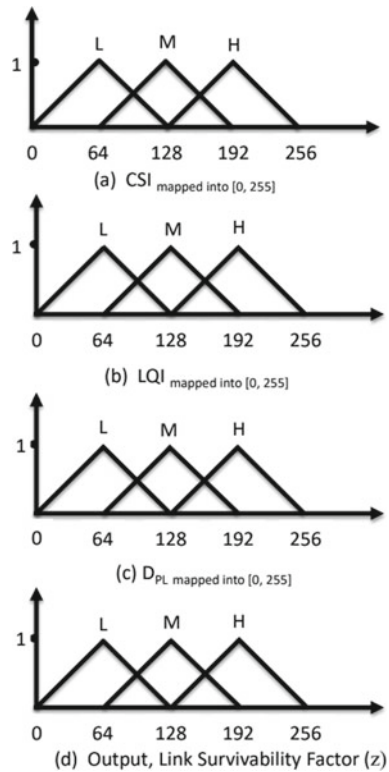
Table 1 Rule base for the fuzzy system

Rules	RSSI	LQI	D_{PL}	Output
1	Low	Low	High	Low
2	Low	Low	Medium	Low
3	Low	Low	Low	Medium
4	Low	Medium	High	Low
5	Low	Medium	Medium	Low
6	Low	Medium	Low	Medium
7	Low	High	High	Low
8	Low	High	Medium	Medium
9	Low	High	Low	Medium
10	Medium	Low	High	Low
11	Medium	Low	Medium	Medium
12	Medium	Low	Low	Medium
13	Medium	Medium	High	Low
14	Medium	Medium	Medium	Medium
15	Medium	Medium	Low	High
16	Medium	High	High	Medium
17	Medium	High	Medium	Medium
18	Medium	High	Low	High
19	High	Low	High	Medium
20	High	Low	Medium	Medium
21	High	Low	Low	High
22	High	Medium	High	Medium
23	High	Medium	Medium	High
24	High	Medium	Low	High
25	High	High	High	Medium
26	High	High	Medium	High
27	High	High	Low	High

In standard CSMA/ CA, back-off exponent (BE) is calculated using binary exponent back-off algorithm in which the value of BE is initially assigned a minimum value, i.e., *macMinBE* that is increased by one unit each time a collision is detected. But in our system, the value of BE is set on the basis of channel characteristics. That is, BE would have a lesser value if the signal strength and the link quality indicator are high and the path loss distance is less for the link between a sender S1 and its receiver R1. If another sender (S2)—receiver (R2) pair has lesser value for RSSI and LQI for the link between them, and higher value for D_{PL}, then they may get allocated with a bigger value for BE. Therefore, the sender S1 first accesses the shared medium over S2. Thus, it makes sure that the node which has better performance

uses the channel. This will lead to the better usage of the medium and hence higher throughput and reliability. In the proposed DCW adaptation, instead of increasing the value of BE each time whenever a collision occurred it will set the value with the output of the fuzzy inference engine. LQI, RSSI, and D_{PL} are the inputs to the fuzzy system, and the corresponding output is based on the fuzzy rules set in the system.

For having commonality between the inputs and also with the output, values of the inputs are mapped into (0, 255). The maximum value of the signal strength is mapped to 255, and the least value of the signal below which the packets are dropped by the receiver is assigned as 0. The generated output of the fuzzy inference engine is also in the range (0, 255). The generated output of the system is mapped back to (macMinBE, macMaxBE). The value of CW is always set to 2 in CSMA/CA standard, which is the number of times the clear channel assessment has been performed.

In the proposed system, the value of CW is also calculated based on the channel characteristics using the fuzzy engine. The node with lesser link quality and signal strength has to do more clear channel assessments than the nodes with better link characteristics. This will also lead to better performance since the nodes with lesser signal values are deferred from sending than the nodes with better characteristics. To avoid the starvation problem for the weak nodes, prioritized sending and GTS slots in the structure of superframe in IEEE 80.15.4 standard can be used. The system with GTS management will be evaluated in the future work, and the present work is not considering the prioritized transmission.

4 Experimental Results and Discussion

Network Simulator-2 is used to carry out the simulations. Transmission range and the carrier sensing range of a node are set to 20 m. Nodes are deployed in a square area of dimension 20 m × 20 m. Figure 3 represents a sample topology. All nodes are pairwise in the transmission range of each other. Results are obtained with the subsequent increase in the number of sender–receiver pairs inside one transmission range in order to analyze the behaving patterns of the protocols in situations of a high number of collisions. Each result is averaged across hundred simulation runs. The proposed fuzzy-based channel allocation protocol with dynamic contention window is abbreviated as DCW, and the standard static contention window adaptation of CSMA/CA of IEEE 802.15.4 MAC is abbreviated as CW in the figures and subsequent sections of the paper.

At the application layer, nodes are participating in the network traffic during particular time intervals. Each source node in the network transmits packets at the beginning of an interval T_i. After this packet transmission period, nodes move to the idle period during which it does not participate in the network traffic. This is a typical application scenario in IoT networks. The nodes which are deployed for sensing the working status of machines, or some other temperature and infrared sensors, etc., will collect the data and aggregate them and then send to their corresponding receivers.
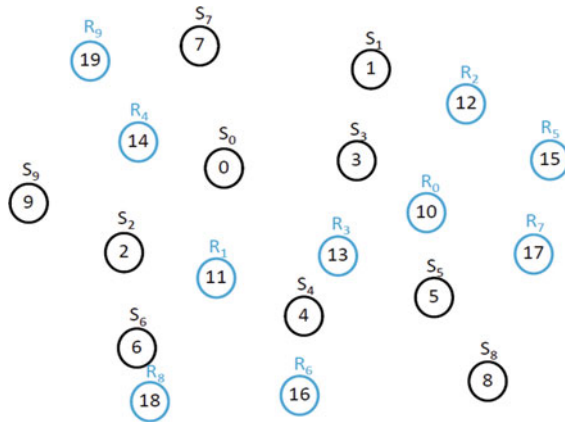
Fig. 3 Sample topology

After this period nodes move to the idle state which consumes less power. In the simulation, protocols are compared by changing the packet interval (T).

Figure 4 evaluates the network throughput. Two methods are compared by calculating average network throughput for all the simulations with different packet periodicities. Our proposed algorithm experiences higher throughput at different traffic levels. Figure 5 displays the comparison of the end-to-end packet delay in the topology. At higher network traffic the proposed algorithm experiences more delay compared to the existing standard. This is because congestion at the nodes would be higher at higher traffic, and hence more back-offs. The range of back-off exponent has to be decided dynamically to resolve this problem, which we will consider in the future work.

Figure 6 shows the number of packets received at the receivers when there are multiple pairs of sender and receiver nodes in the same 20 m transmission range.
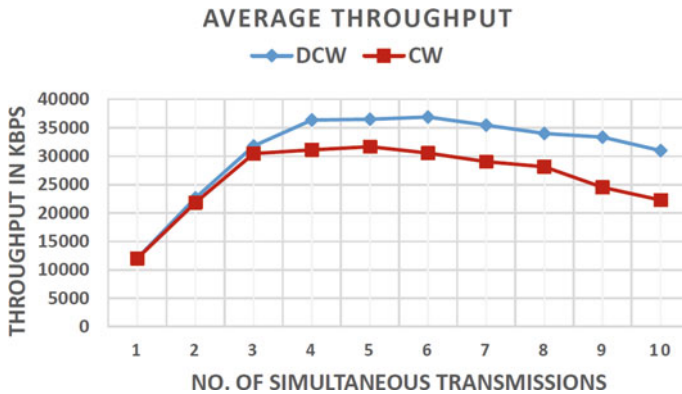
Fig. 4 Network throughput
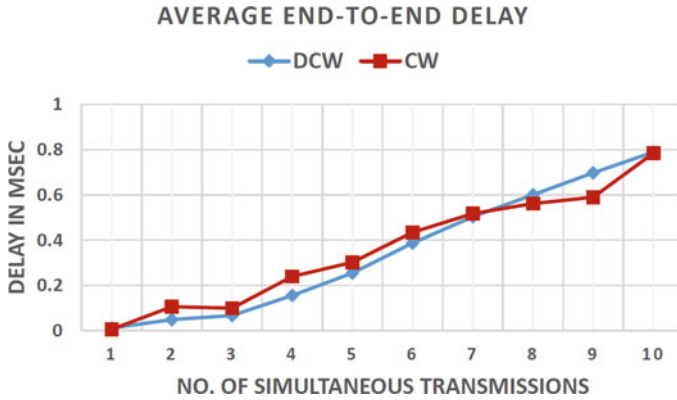
Fig. 5 Average end-to-end delay

The packet interval is being changing during the course of simulation. The time interval between two successive packets is adjusted from 10 to 100 ms. As packet interval decreases, network traffic increases. Both the graphs show that when the value for contention window is evaluated dynamically using fuzzy inference system, the number of packet drop is reduced in the network. This is because, if contention window is changed dynamically, the nodes which have higher link quality would get access to medium and hence can communicate successfully without any interruption. But when the packet interval is very less, there is no significant improvement in the performance. If the packet interval is less, many senders in the network may get the same value for the contention window, and hence, all those nodes would have same back-off counter. They all will get the access to medium at the same time and resulted in a collision. So, when packet periodicity is less, there is no significant decrease in packet drop for proposed mechanism compared to the existing standard.

5 Conclusion

Wireless personal area networks have predominantly become important recently in the domain of networking with its significant applications to serve this world. The nodes in the networks make use of the radio transceivers and wireless communication medium to transmit and receive the data packets, it is likely to interfere each other in high-traffic IoT scenarios like Industrial 4.0 applications. So, the algorithms must be designed to perform in situations that have to deal with extreme traffic conditions with high interference on the network links. The proposed fuzzy-based dynamic channel allocation algorithm will try to allot the shared channel to nodes that are having healthier link quality metrics. Simulation results suggest that the proposed channel allocation mechanism could observe improved network throughput
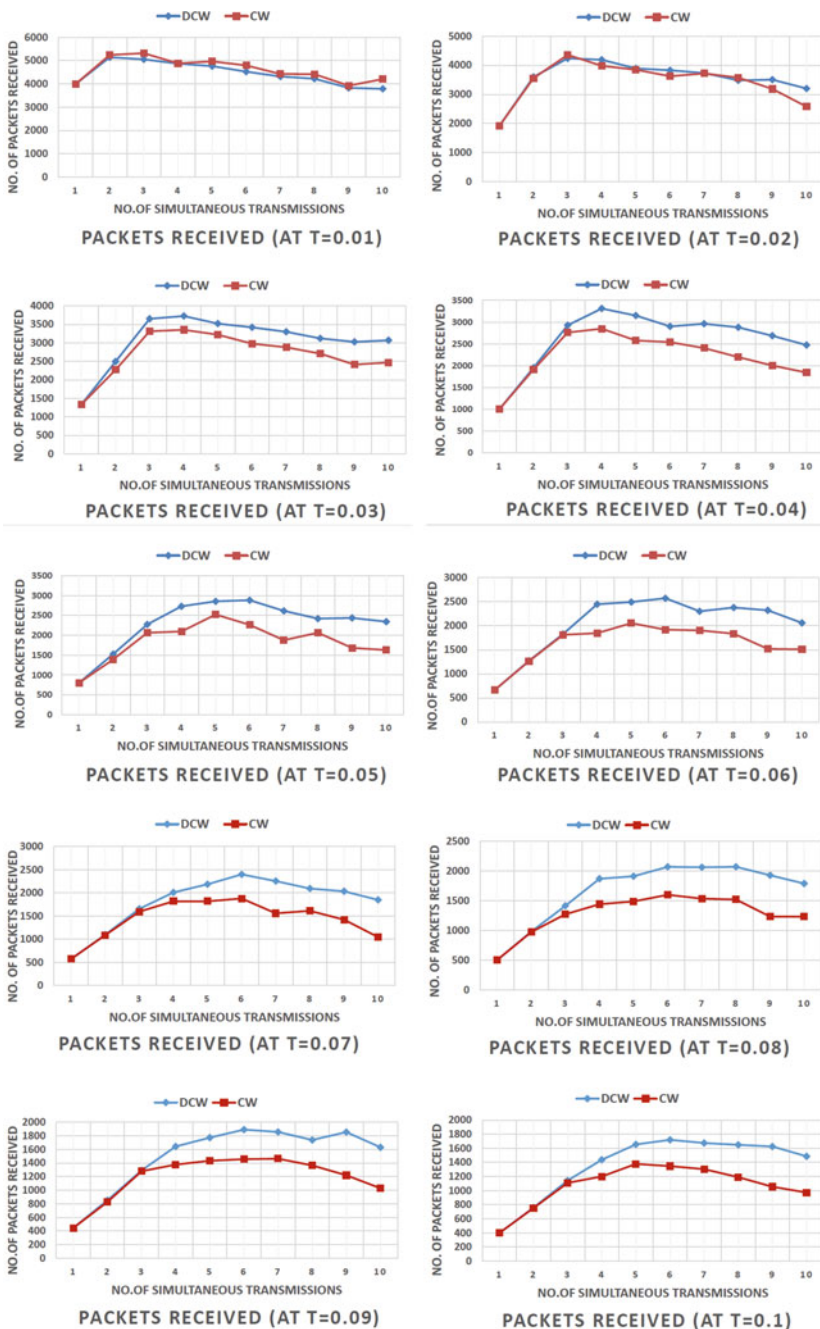
Fig. 6 Number of received packets at the receiver nodes versus number of simultaneous sender–receiver connections in the topology with varying packet interval ($T = 0.01$–0.1 s)

and perform better in terms of packet drop rate compared to the standard MAC procedure. However, at high network traffic, the proposed algorithm experiences a little higher end-to-end delay for the packets. To resolve that the back-off exponent value should also calculate dynamically based on the network scenario. Also, the prioritized transmission with guaranteed time slot management will resolve the starvation problem with the nodes to acquire the medium, and those will be incorporated in the future enhancements.

References

1. Min J, Kim Y, Lee S, Jang TW, Kim I, Song J (2019) The fourth industrial revolution and its impact on occupational health and safety, worker's compensation and labor conditions. Saf Health Work 10(4):400–408
2. ur Rehman MH, Ahmed E, Yaqoob I, Hashem IA, Imran M, Ahmad S (2018) Big data analytics in industrial IoT using a concentric computing model. IEEE Commun Mag 56(2):37–43
3. Begić H, Galić M (2021) A systematic review of construction 4.0 in the context of the BIM 4.0 premise. Buildings 11(8):337
4. Atzori L, Iera A, Morabito G (2017) Understanding the internet of things: definition, potentials, and societal role of a fast evolving paradigm. Ad Hoc Netw 1(56):122–140
5. Liu Y, Tong KF, Qiu X, Liu Y, Ding X (2017) Wireless mesh networks in IoT networks. In: 2017 International workshop on electromagnetics: applications and student innovation competition, May 30. IEEE, pp 183–185
6. Industrial wireless sensor network market size, share & trends analysis report by component (Hardware, Software, Service), by type, by technology, by application, by end use, and segment forecasts, 2019–2025. https://www.grandviewresearch.com/industry-analysis/industrial-wireless-sensor-networks-iwsn-market. Last Accessed 24 Mar 2022
7. IoT market analysis by component (Devices, Connectivity, IT Services, Platforms), by application (Consumer Electronics, Retail, Manufacturing, Transportation, Healthcare) and segment forecasts to 2022. https://www.grandviewresearch.com/industry-analysis/iot-market. Last Accessed 24 Mar 2022
8. Elappila M, Chinara S, Parhi DR (2018) Survivable path routing in WSN for IoT applications. Pervasive Mob Comput 1(43):49–63
9. Elappila M, Chinara S, Parhi DR (2020) Survivability aware channel allocation in WSN for IoT applications. Pervasive Mob Comput 1(61):101107
10. IEEE standard for local and metropolitan area networks—part 15.4: low-rate wireless personal area networks (LR-WPANs), IEEE Std 802.15.4 (2011). (Revision of IEEE Std 802.15.4-2006), pp 1–314
11. Ye Q, Zhuang W (2016) Distributed and adaptive medium access control for internet-of-things-enabled mobile networks. IEEE Internet Things J 4(2):446–460
12. Nekooei SM, Chen G, Rayudu RK (2015) A fuzzy logic based cross-layer mechanism for medium access control in WBAN. In: 2015 IEEE 26th annual international symposium on personal, indoor, and mobile radio communications (PIMRC), Aug 30. IEEE, pp 1094–1099
13. Tushar T (2016) A hybrid signal feature based MAC improvement to Zigbee network optimization. In: 2016 2nd International conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB), pp 282–286
14. Patro A, Chinara S, Elapila M (2017) A dynamic contention MAC protocol for wireless sensor networks. In: Proceedings of the international conference on high performance compilation, computing and communications 2017 Mar 22, pp 97–101
15. Zhou X, Zheng C, Liao M (2015) Full-feedback backoff algorithm for distributed wireless networks. In: 2015 International wireless communications and mobile computing conference (IWCMC), Aug 24. IEEE, pp 1079–1084

16. Chen Z, Lin C, Wen H, Yin H (2007) An analytical model for evaluating IEEE 802.15.4 CSMA/CA protocol in low-rate wireless application. In: 21st International conference on advanced information networking and applications workshops (AINAW'07) 2007 May 21, vol 2. IEEE, pp 899–904
17. Farahani S (2011) ZigBee wireless networks and transceivers. Newnes
18. Chalhoub G, Livolant E, Guitton A, van den Bossche A, Misson M, Val T (2009) Specifications and evaluation of a MAC protocol for a LP-WPAN. Ad Hoc Sens Wirel Netw 7(1–2):69–89
19. Cho J, An S (2009) An adaptive beacon scheduling mechanism using power control in cluster-tree WPANs. Wireless Pers Commun 50(2):143–160
20. Abdulkarem M, Samsudin K, A Rasid MF, Rokhani FZ (2022) Contiki IEEE 802.15.4 MAC layer protocols: implementation and evaluation of node's throughput and power consumption. Wireless Pers Commun 22:1–24
21. Mehta R, Lobiyal DK (2021) Adaptive cross-layer optimization using mimo fuzzy control system in ad-hoc networks. Adhoc Sens Wireless Netw 1:49

Preterm Birth Classification Using KNN Machine Learning Algorithm

K. Naga Narasaiah Goud, K. Madhu Sudhan Reddy, A. Mahesh, and G. Revanth Raju

Abstract Premature births are on the rise all around the world, and there is currently no way to prevent them. The recent study is focused on the examination of ECG records. It includes data on the electrophysiological characteristics of the mother's and foetal heart signals. The purpose of this study is to employ the KNN classifier to categorise foetal ECG heartbeats and predict premature delivery. In this study, 50 ECG signals were collected and preprocessed with the filters NLMS and FIR. FFT was used to extract the function from the preprocessed data. It is uncertain how to classify the signals using the retrieved characteristics. As a result, the classification is carried out using the MATLAB software's Classification Learner programme. By analysing the ECG signals using qualifying criteria, selected features, and target value. ECG signals were classified as either term or preterm.

Keywords Pre-term · ECG signals · FIR · NLMS · FFT · KNN

1 Introduction

Preterm birth happens when the cervix is opened after the fourth month but before the ninth month of the pregnancy, leading to frequent contractions. The earlier a baby is born prematurely, the higher the risk to the health of newborns. Preemies may suffer from long-term physical and mental impairments. Several factors have been linked to an increased chance of preterm delivery, including twin or triplet pregnancy, uterine difficulties, some chronic conditions, such as high blood pressure, and life experiences that are distressing. Preterm birth concerns include birth weight will be low, breathing difficulties, underdeveloped organs, and vision problems. The proportion of preterm newborns has steadily climbed during the previous few decades.

K. N. N. Goud (✉) · K. M. S. Reddy · A. Mahesh · G. R. Raju
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, India
e-mail: knngoud@gmail.com

The amniocentesis maturation process is a test used to determine the lung maturity of a newborn. This diagnosis is also used for the detection of amniotic fluid infections. Magnesium sulphate is a medication used to prolong pregnancy and reduce the risk of embryonic brain damage before 32 weeks of pregnancy. Tocolytics are medications that are used to temporarily cease contractions. Preterm delivery has been classified as slightly premature (32nd–36th week of pregnancy), moderately premature (28th–31st week of pregnancy), and very premature (32nd–36th week of pregnancy) (24th–27th week of pregnancy).

Correcting mullein abnormalities, obtaining periodontal treatment, and having antibiotic medication are all ways to avoid preterm delivery prior to pregnancy. Early avoidance of increased premature risk involves primarily the early detection and medication for depression, as well as the advancement of convertible lengthy contraception, secondary screening for the early detection and tocolytics research and tertiary prevention sought to reduce maternal morbidity and death after diagnosis. The identification of preterm births is a time-consuming process. Despite the fact that certain study publications are displayed, they focus on uterine electrical impulses from the abdominal cavity (with 91% Sensitivity, 84% Specificity, and 12% Mean Error Rate). The goal of this study is to use ECG signals to distinguish between term and preterm delivery. The process's accuracy, sensitivity, and specificity can be improved using this appropriate recommended technique.

2   Literature Overview

(1)   A. Diab, M. Hassan, B. Karlsson, C. Marque

A study of the efficiency and responsiveness to signal intricacy of several nonlinear algorithms consisting of noise and not consisting of any noise injected into artificial signals was performed. Following that, the researchers employed nonlinear methods to analyse a series of uterus electrical surges recorded during pregnancy and labour. According to the findings, temporal alteration is an efficient strategy for identifying pregnancy and labour signals, making it a feasible alternative for genuine, typically noisy signals. From a clinical viewpoint, we will attempt to forecast normal and then preterm labours using these data.

(2)   M. Hassan, C. Muszynski, A. Alexandersson

The purpose of this research is to examine a novel approach for evaluating uterine electromyography's nonlinear correlation (EMG). The use of this approach may enhance pregnancy monitoring, labour detection, and premature labour detection. Uterine EMG signals collected from a four-by-four matrix of electrodes on the individuals' abdomen are employed in this study.

(3) M. Zardoshti, B.C. Wheeler, K. Badie, R. Hashemi

A range of EMG characteristics have been investigated in this article for control of myoelectric upper extremity prosthesis. The robustness, computational cost, and movement class discrimination of these characteristics have been examined for various temporal window widths and noise levels.

(4) A. Phinyomark, P. Nuidod, L. Phukpattaranont

It has been widely employed in the categorization of EMG patterns for both clinical and technical purposes. The purpose of this work was to explore the utility of obtaining EMG features from multiple-level wavelet decomposition and reconstruction. To extract relevant resolution components from the EMG signal, an appropriate wavelet-based function was applied. Throughout the operation, noise and extraneous EMG components were removed. EMG pattern categorization using feature extraction and wavelet transform coefficient reduction.

(5) Malamud Hassan, Jeremy Terrine, Charles Muszynski, Augier Alexanderson, Catherine Marque and Bryn jar Karlsson

The purpose of this research is to examine a novel approach for evaluating uterine electromyography's nonlinear correlation (EMG). The use of this approach may enhance pregnancy monitoring, labour detection, and preterm labour detection. Uterine EMG signals collected from a four-by-four matrix of electrodes on the individuals' abdomen are employed in this study.

3 Existing Methodology

Because preterm delivery is the main cause of newborn mortality, it is critical to identify pregnant women who are at risk of premature labour. Monitoring the electrical activity of uterine muscle appears to be a potential strategy for observing high-quality electrohysterographic data. The created apparatus allowed for the recording of signals via electrodes affixed to the abdomen wall, as well as the derivative of quantitative characteristics defining the detected contractions. Our study included patients with normal pregnancy and with signs of preterm labour. The differentiation between the signals of each patient in each group was categorised using a nonlinear Lagrangian support vector machine. The output of this method distinguishes between the activity of uterine muscles in normal pregnancy and with preterm labour and gives the output as normal or abnormal birth.

SVM

Support vector machines (SVMs) are supervised machine learning algorithms that are both powerful and adaptable. They are used for classification and regression. However, they are most commonly utilised in categorization difficulties. SVMs were initially presented in the 1960s, but they were enhanced around 1990. When

compared to other machine learning algorithms, SVMs have a distinct implementation method. They have recently become immensely popular due to their capacity to handle several continuous and categorical variables.

Disadvantages

- Picking a good kernel function is difficult.
- Long training time for big data sets.
- Difficult to grasp and interpret the final model, varied weights, and individual impact.

4 Proposed Methodology

In this proposed methodology, we use a KNN classifier to predict the preterm and normal delivery. ECG signals are used as input signals to categorise whether the term or preterm birth occurs. From Fig. 1, we can see that the input signals are given to the preprocessing block, the block consists of FIR and NLMS filters to reduce the noise and phase distortion from the signal. And then the output of this block is given to the Feature extraction block where the features of the signals are extracted from the signals which are FFT, mean and standard deviation. And the optimised features were selected from the feature extraction block by the Particle Swarm Optimisation Technique, and these features are given to the KNN classifier. Initially, the KNN classifier is trained with both term and preterm data signals. Then the KNN classifier predicts whether the test signal is a normal or preterm signal with the trained data and gives the output as normal or preterm birth.

The K-nearest neighbours (KNN) method is a type of supervised machine learning technique. The purpose of a supervised machine learning method is to train a function with the formula $f(X) = Y$, where X is the input and Y is the result. KNN may be used for both classification and regression. We shall solely discuss categorization
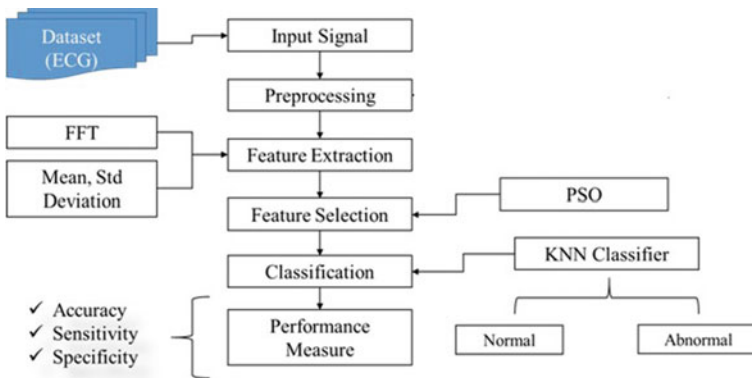
Fig. 1 Block diagram of preterm birth classification

in this post. However, there is only a little difference in regression. Lazy learning implies that the algorithm learns in nearly no time because it just saves the data from the training phase (no learning of a function). The saved data will then be utilised to evaluate a new query point.

A non-parametric approach does not make any assumptions about distributions. As a result, KNN does not need to discover any distribution parameters. During the parametric technique, the model discovers new parameters, which are then utilised for prediction. KNN's single hyperparameter (given by the user to the model) is the number of points that must be evaluated for comparison purposes.

5 Results

Advantages

- The following are the primary benefits of KNN for classification:
- Very easy implementation.
- Flexible in terms of operating space; classes, for example, do not need to be differentiable.
- As additional cases with known classifications are supplied, the classifier may be updated online at a low cost.
- Accuracy is more compared to the SVM (Fig. 2).

6 Conclusion

Finally, we can able to predict preterm or normal birth by using the KNN classification algorithm. In this, we use ECG signals as test signals, these test signals have undergone several filters to remove noise and phase distortion and optimised features were extracted from the test signals and compared by using the KNN classifier and it gives the output as preterm or normal birth. By predicting these preterm births, we can able to stop the number of child deaths at the delivery by taking proper care and medication.
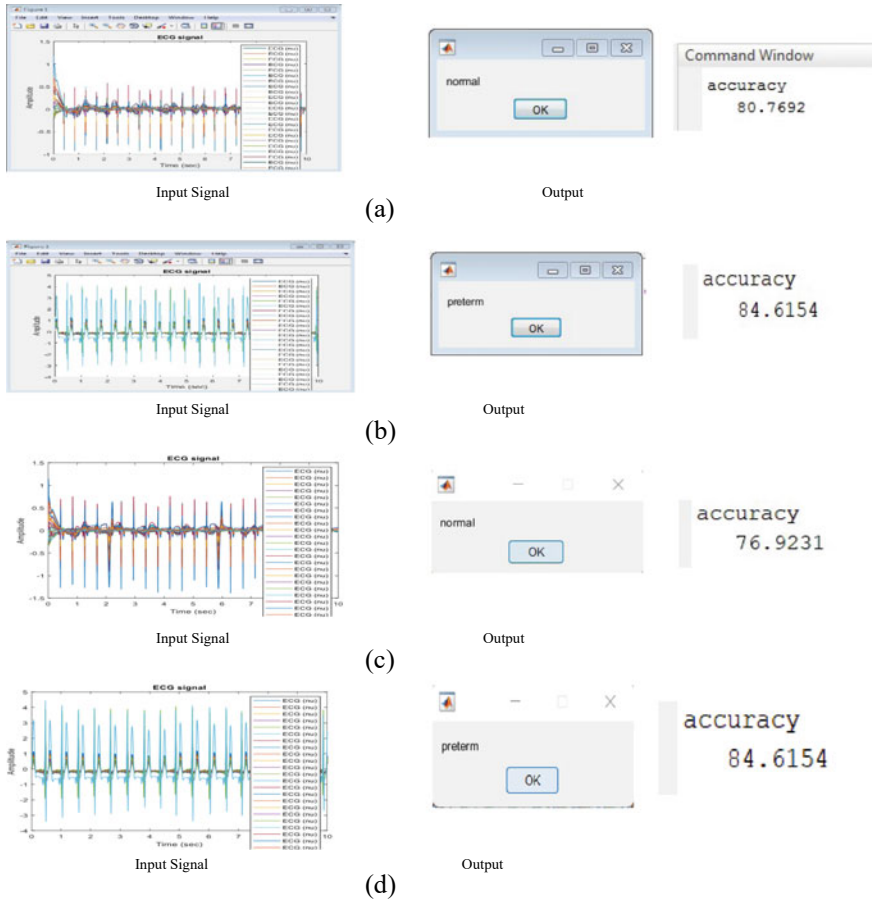
Fig. 2 **a**, **c** Represents normal birth **b**, **d** represents preterm birth

References

1. Diab A, Hassan M, Karlsson B, Marque C (2013) Decimation effect on the classification rate of nonlinear analytical methods applied to uterine EMG signals. Utc.fr, p 12
2. Hassan M, Muszynski C, Alexandersson A (2013) Nonlinear external uterine electromyography association analysis. IEEE Trans Biomed Eng 60(4):1160–1166
3. Zardoshti M, Wheeler BC, Badie K, Hashemi R (2013) Evaluation of EMG features for prosthesis motion control. 15(3):1141–1142
4. Phinyomark A, Nuidod A, Phukpattaranont P (2015) Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification. Electron Electr Eng 6(6):27–32
5. Hassan M, Terrine J, Muszynski C, Alexanderson A, Marque C, Karlsson B (2012) Better pregnancy monitoring using nonlinear correlation analysis of external uterine electromyography. IEEE Trans Biomed Eng
6. Huang DS (2018) The local minima free condition of feed-forward neural networks for outer supervised learning. IEEE Trans Syst Man Cybern 28B(3):477–480

7. Santoso N, Wulandari S (2018) Hybrid support vector machine to preterm birth prediction. (IJEIS) 8(2)
8. Ahadi B, Alavi MH, Khodakarim S, Rahimi F, Kariman N, Khalili M, Safavi N (2016) Using support vector machines in predicting and classifying factors affecting preterm delivery. In: (JPS) Summer, vol 7(No3). ISSN 2008-4978
9. Huang DS (2015) Radial basis probabilistic neural networks: model and application. Int J Pattern Recognit Artif Intell 13(7):1083–1101
10. Sim S, Ryou H, Kim H, Han J, Park K (2014) Evaluation of electrogastrogram feature extraction to classify the preterm and term delivery groups. In: Proceedings of the 15th international conference on biomedical engineering IFMB, pp 675–678
11. Vasak B, Graatsma EM, Hekman Drost E, Eijkemans MJ, van Leeuwen JHS, Visser GH, Jacod BC (2013) Uterine electromyography for identification of first stage labor arrest intermnulliparous women with spontaneous onset of labor. Am J Obstet Gynecol 209(3)
12. Ye-Lin Y, Garcia-Casado J, Prats-Boluda G, Alberola Rubio J, Perales A (2014) Automatic identification of motion artefact sin EHG recording for robust analysis of uterine contractions. Comput Math Methods Med
13. Khail M, Alamedine D, Marque C (2013) Comparison of different EHG feature selection methods for the detection of preterm labor. Comput Math Methods Med

IoT-Based Air Quality Monitoring System with Server Notification

N. Penchalaiah, V. Ramesh Kumar Reddy, S. Ram Mohan, D. Praveen Kumar Reddy, and G. N. V. Tharun Yadav

Abstract The level of pollution has risen across time as a result of a range of factors including fast urbanization, increased automobile use, industrialization, and urbanization, all of which have a negative influence on human happiness by directly affecting the health of individuals who are exposed to it. In order to stay on top of things, I have created a spreadsheet. We will create an IoT-based air pollution monitoring system in which we will monitor the air quality over the internet and send an alarm, i.e., a signal, when the air quality drops or rises above a certain level, i.e., when there is a sufficient amount of harmful gases in the air, such as carbon dioxide, smoke, liquor, benzene, and nitrous oxide. On the LCD, it will show the current air quality and also over on the web, so we all can easily monitor it. You might use your computer or smartphone to monitor the pollution level in this internet-based system.

Keywords Gas sensing devices MQ2 · MQ6 · DHT11 · Server · Arduino mega · GPRS/GSM

1 Introduction

Air pollution is the most serious challenge that every country faces, whether developed or developing. Industrialization and an increase in the number of vehicles have led in the generation of a substantial amount of gaseous pollutants, which has raised health concerns, particularly in developing countries' metropolitan areas. Pollution causes mild allergy symptoms such as burning of the throat, eyes, and nose, as well as more serious difficulties such as bronchitis, cardiovascular illness, pneumonia, lung disease, and aggravated asthma. According to a report, air pollution kills 50,000–100,000 people in the United States each year, with 300,000 in the European and above 3,000,000 worldwide. Air pollution is the primary cause of climate change and poor human health. It has caused rising temperatures such as global warming,

N. Penchalaiah (✉) · V. R. K. Reddy · S. R. Mohan · D. P. K. Reddy · G. N. V. T. Yadav
Department of CSE, AITS, Rajampet, India
e-mail: penchalaiah550@gmail.com

global dimming, excessive rain, drought, storms, acid rain, foggy weather. Because of a lack of proper existence facilities, the living things on Earth and under water are going to experience numerous challenges such as life-changing event.

Working of IoT

The IoT process starts with the gadgets themselves, such as smartphones, smart watches, and electronic appliances like televisions and washers, which connect to the development kit. The following are the four main components of an IoT system:

(1) **Sensors/Devices**. Devices or gadgets are essential components for collecting real-time data from your environments. All of this information could be difficult in some way. It may be a simple temperature monitoring sensor, or it could be something from the video feed. A gadget may contain a variety of sensors that perform functions other than sensing. A transportable gadget, for example, could be a device with several sensors, such as GPS and a camera, but your smartphone is not equipped to detect these objects.

(2) **Connectivity**. The data have been collected then forwarded to a cloud infrastructure. A variety of communication methods should be used to connect the sensors to the cloud. The communication medium includes WAN, Bluetooth, Wi-Fi, satellite networks, mobile phones, and so on.

(3) **Data Processing**. After the data have been collected, it is uploaded to the cloud, where it is analyzed by software. This procedure will consist of just verifying the temperature with readings from gadgets such as air-conditioning systems and heaters. However, some tasks, such as detecting objects with computer vision on video, can be extremely difficult.

(4) **User Interface**. The data should be accessible to the end-user in some way, for as by sending an email or text, or by setting up notifications on their phones. The user may require an interface in order to check data on their IoT system. The user, for example, has a camera installed in his home. He wants to see the recordings, and everyone pulls off of a web.

2 Related Work

Zheng et al. [1] developed a new way for implementing an air quality monitoring system using cutting-edge Internet of Things (IoT) technologies. Portable sensors capture air quality data in real time and send it via a low-power wide area network in this system. The IoT cloud processes and analyzes all air quality data. The completed air quality monitoring system, which includes both hardware and software, has been successfully designed and deployed in urban settings. Experiments show that the suggested system is reliable in sensing air quality, and that it can assist disclose air quality change patterns to some extent.

Marinov et al. [2] proposed a modular sensing system with integrated response measures and infrared gas sensors for cost-effective assessment of relevant environmental parameters. The device was put to the test in the city, and the results were

compared to data from local environmental control agency stations. The preliminary findings indicate that this method can be utilized as a cost-effective alternative to professional-grade equipment.

Jha et al. [3] developed a new sensor nodes implementation for comprehensive urban microclimate observation and environment modeling in order to adapt future urban infrastructure. They demonstrated a technology for monitoring, modeling, and eventually modifying the urban microclimate to make additional urban infrastructure more adaptive. They calibrate a variety of sensors (temp, moisture, thermal image, and air sensors), then deploy many sensor nodes throughout the city to collect and feed data. The node does have its own connection system which allows it to connect via the internet with other nodes. Internet of Things concepts are so incorporated. They investigated a variety of sensor systems and network architectures to aid microclimate modeling through the integration of captured and analyzed data.

Gupta et al. [3] presented a model for an atmospheric air pollution monitoring system to detect the concentrations of massive air pollutant gases. Reduced air monitoring nodes with semiconductor gas sensors and Wi-Fi modules are used in the system. This system detects gas concentrations such as CO, CO_2, No_2, and sulfur dioxide using semiconductor sensors.

The data collected by sensors are visualized on a Raspberry Pi 3-based Web Server. A MEAN stack is used to display data on websites. Air quality data were obtained from various cities in South Africa by Chiwewe, Tapiwa, and Jeoffrey Ditsela [4] Ground-level ozone prediction models were created using a machine learning technique applied to the data.

Maslyiak [5] suggested a model environmental monitoring that considers the design of software and hardware systems. Its primary responsibilities are known. This website allows for the forecast of air pollution caused by harmful emissions from motor vehicles, as well as the visualization of modeling findings. Differential equations are how mathematical models are represented. They depict the dynamics and distribution of harmful motor vehicle emissions in terms of both time and space.

Transactions of the IEEE (2019) Abdullah J. Air pollution Monitoring System with Arduino. The goal of this study is to offer a design for a system that will alert occupants to VOC concentration levels in both indoor and outdoor situations. The Arduino-based system will track and analyze total volatile organic compounds (TVOC) before informing the user of their amounts via a wireless communication system so that they can take action.

Transactions of the IEEE (2020) Jha, Rohan Kumar IoT-based air quality monitoring and reporting system. Finally, an Android application that pulls data from ThingSpeak presents the PPM readings as well as the air level of quality of gases. The current model has been successfully implemented and can be used to create real-world systems.

Transactions of the IEEE (2019) Air Quality Monitoring at Hocine Mokrani Using IoT: A Survey. This study attempts to address these needs by analyzing existing research on IoT-based air quality monitoring, with an emphasis on recent developments and issues.

All the devices are initialized, LCD, GSM/GPRS, Gas Sensors, Humidity and

The sensor value are monitored and uploaded to cloud via GPRS

Gases values can be read by public at any time using internet
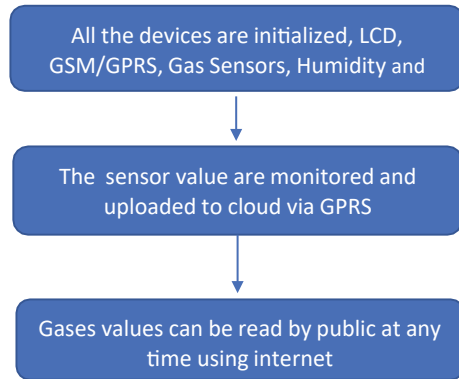
Transactions of the IEEE (2020) Kumar Ajitesh IoT-based Air Quality Monitoring System Design and Analysis. The system can detect local air pollution and create analytical data, which it uses to inform people via a buzzer device implemented. The system technology is so user-friendly and simple to operate that it can be implemented in homes and small spaces.

3   Proposed System

We employed a cloud server in this proposed system as shown in Fig. 1 to record air values that can be monitored from any location using a mobile or PC nearby. A URL will be provided for those values, which we may monitor at any moment. The values of sensors are shown on an LCD.

MQ6 and MQ2 are used to collect various harmful gases such as NH3, benzene, CO as shown in the model of Fig. 2. The sensed data are delivered to the Arduino UNO, where we will display the values on the LCD and send the data to the cloud over ThingSpeak using GSM/GPRS. If the detected value exceeds a certain threshold, a warning could be sent via SMS utilizing GSM/GPRS.

4   System Architecture

Hardware requirements and software requirements combination is a system architecture. The architecture is the interfacing of sensors, Arduino UNO and GSM/GPRS module with computer collected data transferred to cloud through internet.

MQ6 Sensor: Gases such as LPG and butane can be detected or measured with the MQ6 gas sensor. The MQ6 sensor system incorporates a digital pin that allows it to function without the assistance of a microcontroller, which is handy when only one

Fig. 2 Working model flow
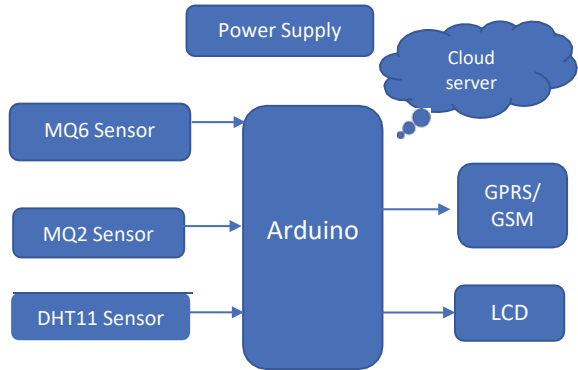
gas needs to be detected. The analog pin must be used for detecting gas concentrations in ppm.

Arduino UNO: It is an Arduino.cc microcontroller board, an accessible electronic platform based primarily on the microcontroller board ATmega328. For connecting to external electronic equipment, the Arduino UNO now offers a USB interface, 6 input analog pins, and 14 I/O digital ports. Six ports of the 14 I/O interfaces can be utilized for PWM output.

MQ2 Gas Sensor: It is being used to indicate the existence of LPG, propane, and hydrogen in the atmosphere. It can also detect methane as well as other flammable steam. It is inexpensive and can be used for a variety of purposes. This sensor is flammable gas and smoky sensitive. The smoke sensor is powered by 5 V. The voltage generated by a smoke sensor shows the presence of smoke; the more smoke, the higher the output. The sensitivity can be adjusted with a potentiometer.

DHT11 Sensor: The device is a digital temperature and moisture sensor with a modest price tag. To detect temperature and relative humidity in real time, simply attach this sensor to any microcontroller (Raspberry Pi, Arduino UNO, etc.). There are two versions of the DHT11 humidity and temperature sensor: a sensor and a module. This sensor is distinguished from the module by pull-up resistors as well as a power-on LED. The DHT11 is a relative humidity sensor. To monitor the ambient air, this sensor uses a thermistor and a capacitive moisture sensor.

GSM/GPRS Module: GSM/GPRS packages are one of the most widely used communication modules in embedded systems. A GSM/GPRS module is a device that connects a microcontroller (or a processor) to a GSM/GPRS network. The General Packet Radio Service (GPRS) and the Global System for Mobile Communication (GSM) are two acronyms for the universal mobile telecommunication and the GSM service, respectively.

LCD: The term "liquid crystal display" is an acronym for "LCD". It is a form of electronic display unit that is found in a variety of circuits and devices, including phones, calculators, computers, and television sets. The most popular displays are

multi-segment illumination diodes and seven-segment displays. The low cost, ease of scripting, animations, and the reality that there are no constraints on displaying different characteristics, special and sometimes even animations are the main benefits of using this module.

Arduino IDE: Arduino Integrated Development Environment (IDE) is a cross-platform development environment for Windows, Mac OS X, and Linux that has been built in C and C++ functions. It is used to transfer files between computers to Arduino-compatible boards and other manufacturer organizations and local that have third-party cores.

ThingSpeak: ThingSpeak is an internet online analytical processing tool that allows the user to collect, visualize, and analyze live data streams. ThingSpeak visualizes data sent to the internet by IoT devices in real time. You can analyze and handle data in real time only with desire to run MATLAB script in ThingSpeak as shown in Fig. 3.

Fritzing: Framework is an open project that seeks to create beginner or hobby Cad model for the creation of circuitry in designed to help designers and creators who are ready to move beyond prototyping and into the creation of a more permanent circuit. The software, inspired by the processing programming language and the Arduino microcontroller, allows a designer, artist, developer, or enthusiast to describe as well as generate a PCB layout for 62 productions. On the linked website, users can share and discuss drafts and experiences, as well as cut production costs. The daywise comparison of matrices is provided in Table 1 and Fig. 4.

The process flow of the entire experimental research is shown in Fig. 5 which initiates with the interfacing with the Arduino UNO board and later processes the data of sensors. The main condition of the process flow relies on the threshold limit, otherwise the data can be stored in the cloud server.

Fig. 3 ThingSpeak module interface
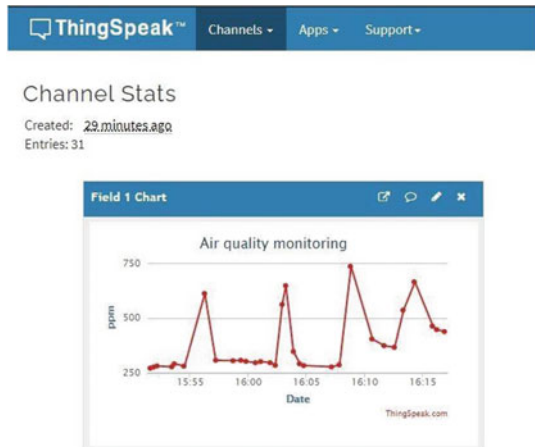
Table 1 Daywise schedule of the metric comparison

Day	CO2 (PPM)	CO (PPM)	Butane (PPM)	Temp
Monday	688	256	0.0	36
Tuesday	679	253	0.0	36
Wednesday	671	250	0.0	36
Thursday	683	249	0.0	36
Friday	668	247	0.0	36
Saturday	663	246	0.0	36
Sunday	654	243	0.0	36
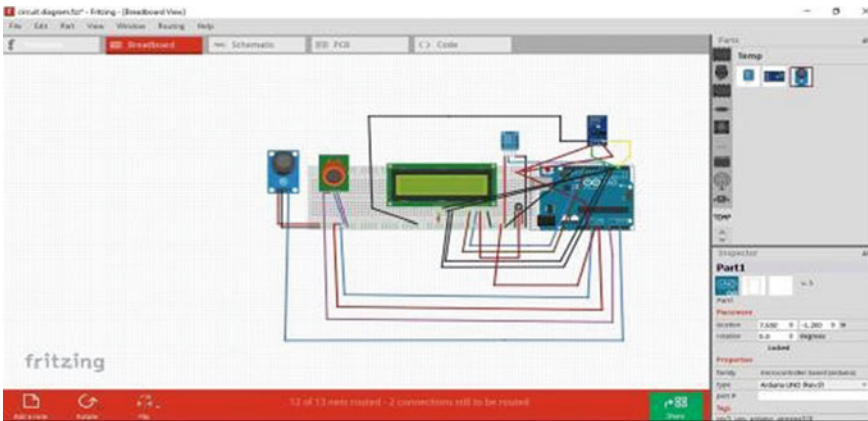
Fig. 4 Fritzing framework

5 Results

The implementation requires use of an Arduino as shown in Fig. 6 which is a global communication system for mobile devices. Our system is self-system on chip (SOC) that provides direct connections to your Wi-Fi network via SIM to any microcontroller. The MQ2 gas sensor can measure LPG, liquor, propane, hydrogen, CO, and even methane. The MQ6 gas device monitors CO levels in the environment ranging from 20 to 2000 ppm and also detecting CO using a cycle of high and low temperatures and detecting CO when the temperature is too low (heated by 1.5 V). The impedance of these sensors increases as the gas concentration increases. The data are transferred to a cloud service, where users can learn about the gas concentrations, temperature, and moisture in the atmosphere. The air quality monitoring systems via ThingSpeak interfaces are shown for CO_2 and CO in Figs. 7 and 8, respectively.
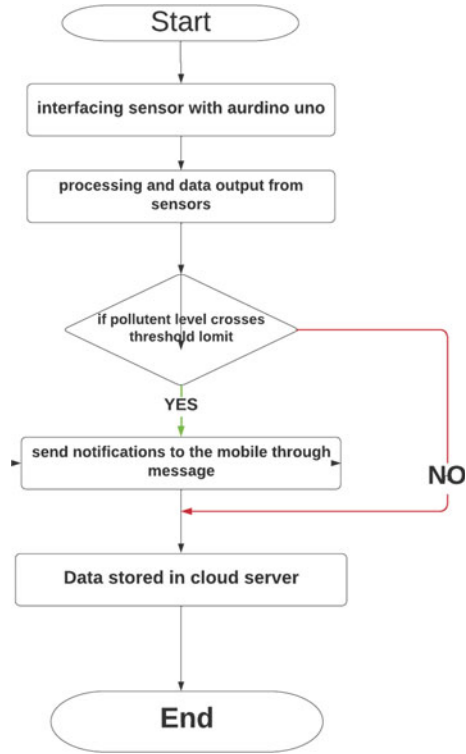
Fig. 5 Process flow of the work

Fig. 6 Prototype module

Fig. 7 CO_2 air quality monitoring
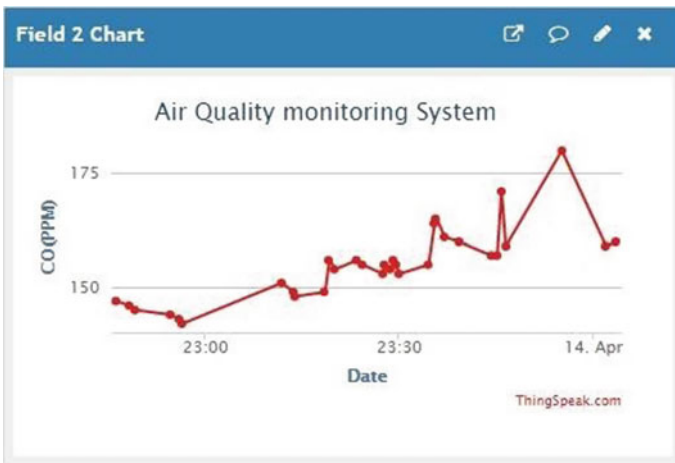
Fig. 8 CO air quality monitoring

6 Conclusion

This system incorporates sensors for detecting pollution-causing parameters. Carbon dioxide sensors are being used. When the amount of these parameters rises, the sensor detects the situation and sends out a warning or indication. The notice appears on the LCD screen. The use of a Pic microcontroller to evaluate the air quality in the neighborhood is offered as a way to enhance air quality. The use of MQ6 and MQ2 sensing applications detects various types of hazardous gases, and Arduino seems to

be at the heart of this proposal, adjusting the entire solution that provides visuals via an LCD.

References

1. Kamaruzzaman SN, Sabrani NA (2011) The effect of indoor air quality (IAQ) towards occupants' psychological performance in office buildings. J Des Built Environ 4(2001):49–61
2. An Office Building Occupants Guide to Indoor Air Quality. E. P. Agency, 1997. [Online]. http://www.epa.gov/iaq/pubs/occupgd.html is a good place to start
3. Erdmann CA, Steiner KC, Apte MG (2002) Indoor carbon dioxide concentrations and sick building syndrome symptoms in the base study revisited: analyses of the 100 building dataset, pp 443–448
4. Sharma A, Sohi BS, Chandra S (2019) SN based forest fire detection and early warning system. Int. J. Innovative Technol Exploring Eng (IJITEE) 8(9):209–214
5. Rao Jaladi A, Khithani K, Pawar P, Malvi K, Sahoo G (2017) Environmental monitoring using wireless sensor networks (WSN) based on IOT. Int Res J Eng Technol (IRJET)
6. Apte MG, Fisk WJ, Daisey JM (2000) Indoor carbon dioxide concentrations and SBS in office workers. Healthy Buildings 1:133–138
7. De Vito S, Fattoruso G, Liguoro R, Oliviero A, Massera E, Sansone C, Casola V, Di Francia G (2011) Cooperative 3D air quality assessment with wireless chemical sensing networks. Proc Eng 25:84–87
8. Tahseenul Hasan M, Chourasia VS, Asutkar SM (2019) A forecasting tool for air quality monitoring built on cloud and IoT. Int J Innovative Technol Explor Eng (IJITEE) 8(10):3821–3832
9. Rai A, (Nigam) Saxena V, Kusrey S (2017) Zigbee based air pollution monitoring and control system using WSN. SSRG Int J Electron Commun Eng 4(6):7–12
10. Shirsath SM, Waghile NB (2018) IoT Based smart environmental monitoring using wireless sensor network. Int. J. Adv. Res. Electr Electron Instrum Eng 7(6):3023–3031

Hand Gesture Recognition Using CNN

N. Penchalaiah, V. Bindhu Reddy, R. Harsha Vardhan Reddy, Akhileswari, and N. Anand Raj

Abstract Computer is useful and important in our life and used in different fields. Interaction between computer and human is accomplished with computer input devices like mouse, printer, keyboard, etc. Using hand gestures as useful medium between human and computer can make the connection easier for disable persons. Disable persons do not have an interaction with normal people because, the normal persons do not know the sign language. We use the CNN for the recognition of hand gestures. As CNN can intake image processing, we use this technique. Recent research has proved the supremacy of convolutional neural network. Whatever, the images that can be captured are compared with datasets and compared the accuracy and give the message of respected dataset. With this software, the disable people can communicate with others. The model of augmented data received accuracy 97.12% which is nearly 4% higher than the model without augmentation (92.87%). So, nonlinearity exists in this type of problem. Also, we added some applications such as mouse movement and volume changes.

Keywords Convolutional neural network · Static hand gestures recognition · Data augmentation

N. Penchalaiah (✉) · V. B. Reddy · R. H. V. Reddy · Akhileswari · N. A. Raj
Department of CSE, AITS, Rajampet, India
e-mail: npc@aitsrajampet.ac.in

1   Introduction

The meaning of gesture has been defined by Bobick and Wilson. According to them, the movement of the body that is intended to gesture is a way of communicating with other agents [1]. In a world of technological advancements, almost 300 million are deaf and 1 millionaire dumb people. For conversations to be possible, knowledge on ways of expression and insights of their standard versions in practice are to be studied, and as a concern toward making their lives better, many research works have been under progress. The following proposed project helps to develop an integrated system which can be useful for deaf/dumb people to easily communicate with normal people. This can be further developed into an innovative communication system which can support mobile or wireless communication for deaf and dumb in a compact device.

The purpose of this study is to explain how to identify hand movements. The main issue is teaching a computer to recognize hand motions. The movement of the fingers and the shape of the hands vary in different hand motions. So, variances are cause of the main specifications of hand gestures that have to be worked. This technique can be applied using the huge data and content information with the images. The process contains of two tasks: feature extraction and the classification. Before recognition of every image must have qualities that are representative of any hand gesture is extracted. Then, classification method should be applied. Here, the main grievance is the extraction of those features and input those features for classification.

Huge features are compulsory for classifying and recognizing hand gestures. Conventional models for pattern recognition could not process original data in original form [2]. Therefore, more efforts are required to extract features from raw data, and they are not automated. The basic approach of the idea involves conversion of one mode of communication to the other which is sign language to text/speech and control mouse using hand gestures. Sign language is a well-structured non-verbal communication skill through which a speaker's thoughts can be meaning fully conveyed wherein each gesture, including movement of head and other body parts, has a meaning assigned to it. In the proposed system, a gesture or sign image is sent to the system which is then evaluated using neural network models like (CNN) convolutional neural networks. It involves different layers where feature extraction and classification steps are performed to enhance hand features extracted from the image. When the input image matches with the given dataset, the output gesture is recognized, and based on that, mouse functions are performed.

Therefore, this paper describes about the recognition of hand movements by developing a CNN model that can evaluate large amounts of picture data and distinguish different types of hand gestures.

2 Related Work

Various works on hand gesture identification using CNN have been done, and only a small amount of noteworthy research in the field has been cited.

A hand gesture recognition system that employs an artificial neural network is evaluated using the form-fitting approach [3]. After filtering, this technique was used to detect hand signs using a color-based segmentation approach on YbrCr. Hand morphology was then used to determine this hand form. The hand and finger motions were received by the ANN. Then, I had a 94.05% accuracy rate with this method.

Gesture movement identification by using an ANN has been identified [4]. In this method, images are separated by skin colors. Selected movements for ANN are converted to pixels using cross sections, boundaries, and their scalar information like edge ratio and aspect ratio. Then, evaluating the feature vectors was sent to ANN for process of training. Received accuracy was 98%.

The technique using Haar-like characteristics to identify hand gestures had spotted [5]. From above results, they used the technique known as AdaBoost algorithm. The total work is spitted into two levels. The top one contains the context-free grammar which is used to recognize movements of hand. In bottom level, movements were identified. Usual string was generated for each input according to grammar. The likelihood of each rule was computed, and the rule with the highest probability for the terminal string supplied was chosen. The rule's hand motion was mistakenly regarded as an input gesture.

Extraction of human characteristics approach has a few minor flaws. This extraction process was difficult, and it is probable that not every available characteristic was extracted correctly. The extraction, on the other hand, was becoming a human wish. In engineering, there is a new automated feature technique that is neither difficult nor tiresome, nor human biased. CNN will extract useful features and gestures from structured data. As a result, workers began to believe and migrate to automated feature techniques in engineering, and deep learning, or CNN, began to evolve.

The algorithm technique which is used to identify hand gestures using the 3D and that is CNN had proposed [6]. The current technique's depth and intensity of hand images are challenging the basis of identification or recognition.

Another technique to identify gesture using the CNN is robust under five features: scale, rotation, translation, illumination, noise, and background were proposed [7].

3 Experimental Methodology

This part gives the description of the dataset, and configuration of CNN was used. The methodology flowchart is shown in Fig. 1. Data collection, pre-processing, CNN configuration, and model construction are all part of the procedure.
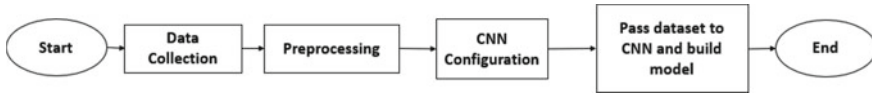
Fig. 1 System framework

Inputting Data and Training Data

Images which are used for our algorithm are captured by webcam and sent to the training. The gestures are done by ten members in front of device webcam. Whatever, the images captured are that must be right hand, and palm should be in front of webcam so that the background complexity decreases. The important one is hand should be vertical and images must be uniform in nature.

A. *Pre-processing*

This technique had applied on the datasets to decrease analytical complication while still receiving high efficiency. First, the backgrounds for each image were deleted using a method that is background subtraction method ZivKovic [8, 9] proposed it. This background reduction is largely based on the K-Gaussian distribution approach, which finds a selective-Gaussian distribution for each and every pixel and provides strong adaptability on diverse situations due to changes in light. After reducing background, the image is obtained.

Then, obtained images are converted into the grayscale images. However, grayscale images obtained contain only one and only color channel which makes easier to learn CNN [10], so that, morphological erosion technique had been applied [11]. Then, filter of median is applied in signal processing which reduces noises [12]. Figure 2 shows the steps of pre-processing. The images are converted and sent to CNN, then resize the image to 50 × 50 pixels.

Fig. 2 Pre-processing working
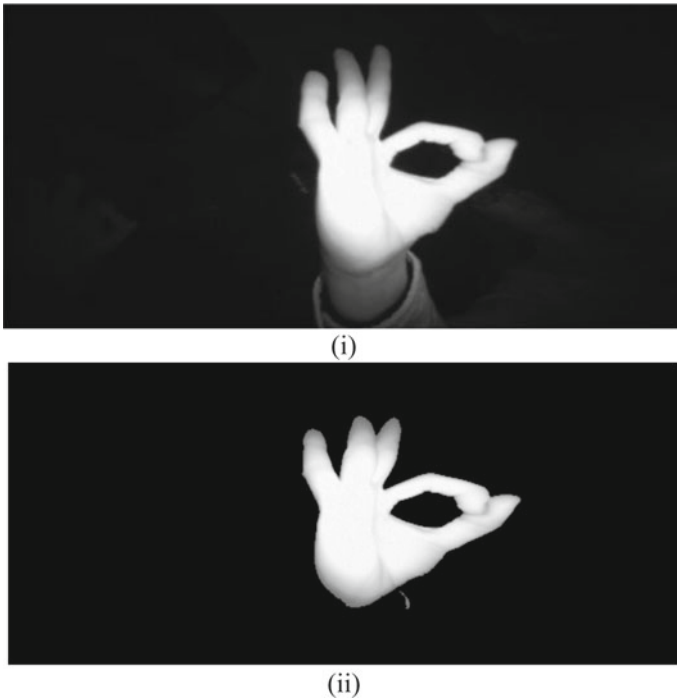
Fig. 3 **i** Captured image **ii** after selecting the largest object

In addition to our own dataset, the project also makes use of another dataset called "Hand Gesture Recognition Database" [13]. Different objects from collected photographs are deleted in this method by computing the largest object, the hand. The effects of gathering larger objects are seen in Fig. 3.

B. *Dataset*

We had selected ten hand gestures of our own which can be saved as model named epoch in CNN environment to recognize datasets for training. Some of images we used are in Fig. 4.

There are ten classes in "Hand Gesture Recognition Using CNN Database" [13]. (Palm, I, Fist, Fist Moved, Thumb, Index, OK, Palm Moved, C, Down). Figure 5 shows an example from the database.

C. *Configuration of CNN*

The CNN which is used in this research is to identify hand movements and is worked on two convolutional layers, two maximum pooling layers, two fully attached layers, and return layer.

Here, three dropout performances are used in network for the prevention of over-fitting [14].
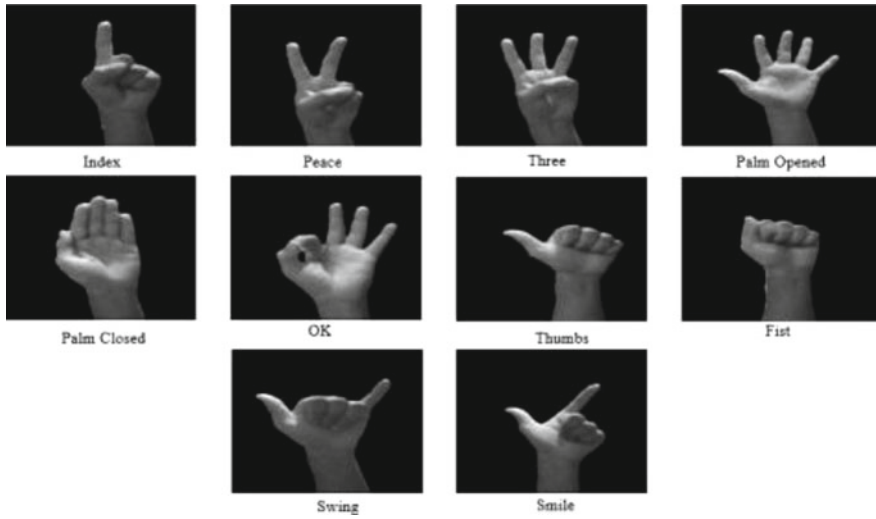
Fig. 4 Self-images datasets

Fig. 5 Some of the images from database of hand

This layer has 64 different filters, each with a 3 × 3 extracted kernel size. The activation function used here is rectified linear unit shortly known as ReLU. This was applicable to give constant [15], and it showed that ReLU performs well compared with other innovation methods such as tanh or sigmoid. It has been input layer; we have to enter input size. The stride is settled to original. Input shape is identified 50

$\times$ 50 $\times$ 1 that meant the grayscale size image of 50 $\times$ 50 could be applicable to the network. Presently, this layer gives the function maps and passes next to the layer.

The CNN features a maximum pooling layer with a 2 $\times$ 2 pool size that takes the maximum number of values from the 2 $\times$ 2 window size. The empty size of our representation is gradually reduced until just the maximum value is stored in the pooling layer, which then deletes the remainder. Because it only chooses the most important information, this layer aids the network's understanding of the pictures.

Another CNN with many types of filters, a 3 $\times$ 3 kernel size, and original pace is the next layer to consider. ReLU was used as the activation function once more in this layer. A maximum pooling layer with a pooling size of 2 $\times$ 2 follows this layer. In the current layer, the first released one was mixed, which typically discards 25% of total neurons to prevent model over-fitting. The current layer's output has been sent to the flatten layer. The flattened layer achieves the output that we acquired from the previous layers, and these are flattening to vector from a 2D matrix. This layer currently only permits fully linked layers to process the data that has been received up to this point.

The other layer is starting a fully linked layer with 256 nodes, with ReLU as the activation function up to this point. The next layer is the dropout layer, which removes 25% of the neurons to prevent over-fitting.

The following fully attached layer, which employs ReLU as an activation layer, contains 256 nodes to reach the created vector supplied by the first fully connected layer. This is followed by a dropout layer that removes 25% of the neurons to avoid over-fitting.

Returned output layer contains ten nodes specifying to each and every class of hand gestures. The current layer employs the softmax method [16] as an activation method and gives the value of probability for each class.

Present model is now the compiled using Stochastic Gradient Descent (SGD) [17] method containing learn rate of 0.001. Since the present model is checked for two or more classes, the evaluation loss and the categorical cross-entropy approach [16] were utilized. Finally, the loss evaluation and accuracy were determined in order to create a track on the computation procedure.

This type of configuration was choosing after trying different combination of nodes and neural layers.

D. *Implementation of System*

(1) *Base Dataset Training*: The model of pre-processing was achieved after training using the base dataset.

(2) *Expanded Dataset Training*: The dataset augmentation had done. Data augmentation is a technique to increase the amount of data by initializing zoom, rotation, flip, and so on [18]. Present process improves the data, and CNN needs to know the differences in images in the dataset. The effect of augmentation is clearly shown in Fig. 6. To obtain the demonstration of accuracy, any random image was used.
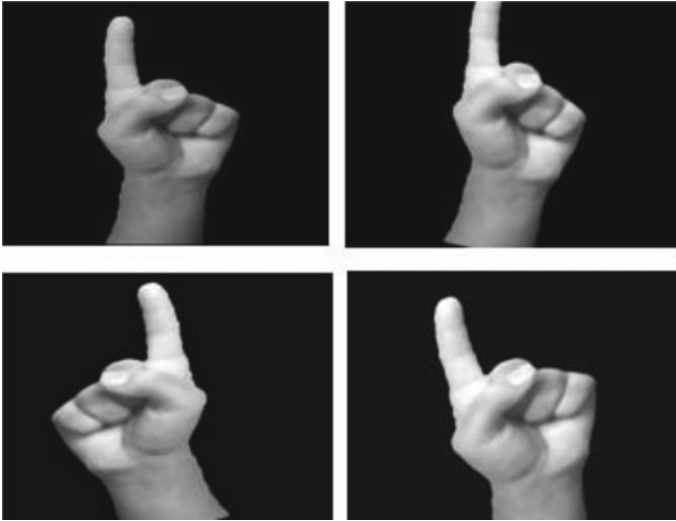
Fig. 6 Data augmentation effects

To work on this system, the programming language Python was used, and known IDE of Python Spyder was worked to run code of this project. Most common library known as Keras was used in building this CNN classifier. For pre-processing, we used PHIL library. Sklearn was used for evaluation of the confusion matrix. We used matplotlib library in the CNN classifier for visualizing model accuracy. NumPy module had been used for the techniques of array.

The dataset working process is divided into two stages.

4 Results

Present part explains the consequences achieved from the project using convolutional neural network algorithm configuration. The experimental output shows model obtained which had augmented with artificial data received 97.12% accuracy compared to before data which is 4% higher than any augmented data.

The accuracy graph is shown in Fig. 7. It gives the progress of an accuracy of model augmented that is better than a non-augmented model which means invariant model.

Confusion matrices with two models were given in Fig. 8. Non-bounded diagonal values represent the number of distinct tuples, whereas diagonal values represent the number of tuples correctly classified by two models. Higher diagonal value gives good performance. The matrices say non-augmented model could not perform well on three classes (Index, Peace, Three), and second augmented model shows a best
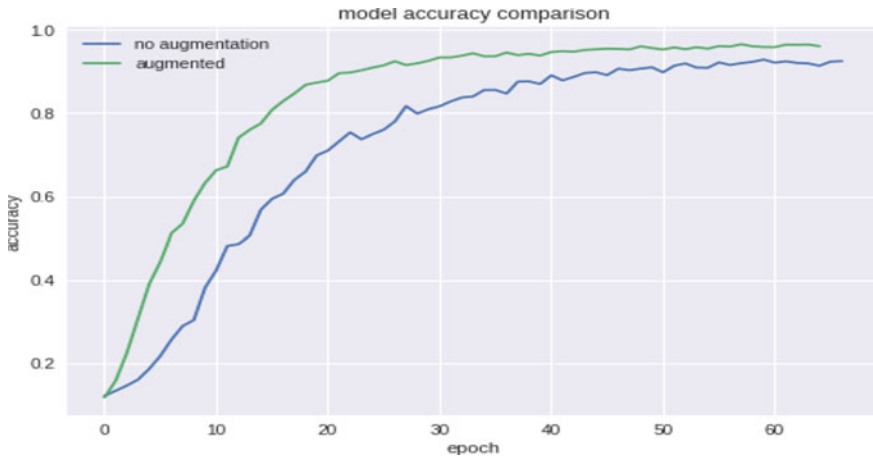
Fig. 7 Accuracy

performance with three classes because model had provided more amount of improvised data. So that performance was given good experience in the cases of augmented model.

Based on Fig. 9, the accuracy of the training and testing information was near in each epoch, indicating that the model of data augmentations was not over-fitted.

The same experiment was repeated using a 65:35 split in training and testing. The accuracy for the 65–35 split was found to be 96.57%, indicating that the model's adaptability is a large test set.

The equal-sized dataset was given into support vector machine as input which is shortly known as SVM and *K*-nearest neighbors which is shortly known as KNN models. Using this model's received accuracy was of 72–75%. The reason of low performance shown by support vector machine and *K*-nearest neighbor is one of the adaptability issues without a linear dataset. Dataset was provided in original format, so the accuracy is lesser than the CNN based.

The accuracy of the proposed technique was 98.95% when it was applied to the "Hand Gesture Recognition Database" [13] being 70:30 splitting ratio. This indicates that the proposed method is adaptable to various datasets. The confusion matrix derived from dataset [13] is shown in Fig. 10.

5 Conclusion and Future Work

This research explores the communication is an important part of our lives. Deaf and dumb people being unable to speak and listen, I have a lot of issues talking with regular folks. People with disabilities can help themselves in a variety of ways. These disabilities try to communicate. The use of sign language, or hand gestures, is one
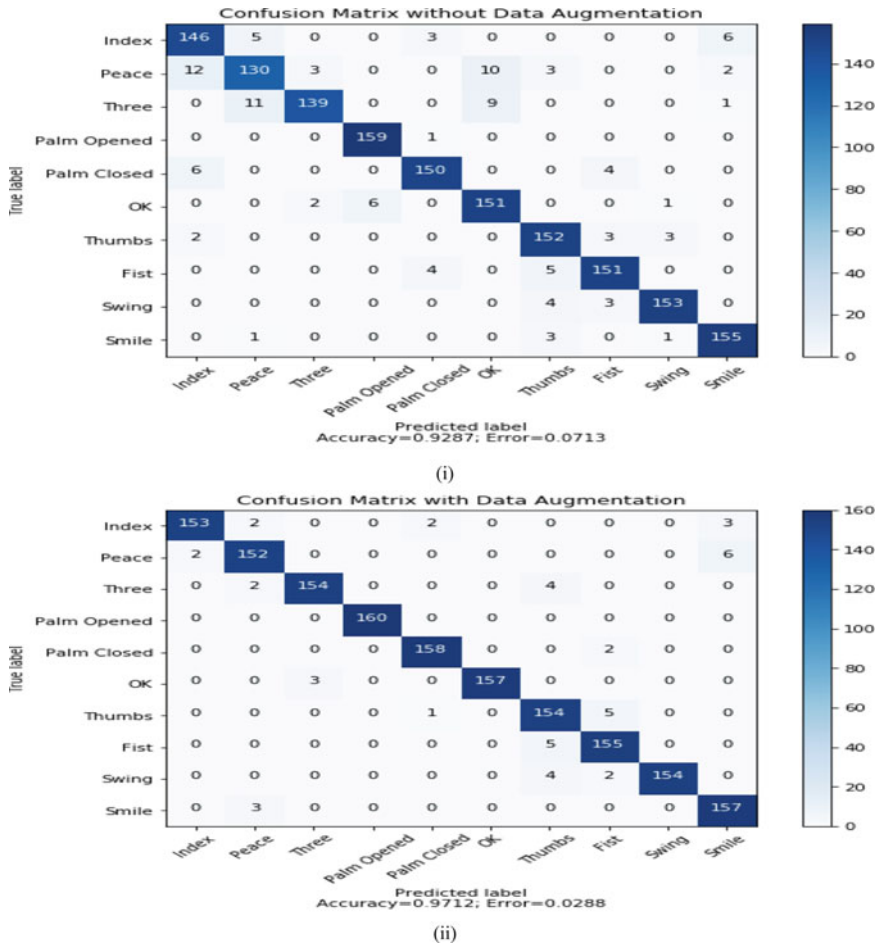
Fig. 8 Confusion matrix **i** non-augmented model and **ii** augmented models

of the most used methods. It is important to create a gesture recognition application of actions of sign language so that deaf and dumb and others who have disabilities can understand sign language and can converse effortlessly.

The goal of this project is to take a first step toward breaking the barrier in communication between the normal people and deaf and dumb people with the help of sign language. The accuracy of the model obtained was appreciable when convolutional neural network was used. Although the system only recognizes hand gestures perfectly, better implementation is still possible in real time. Example, by giving the knowledge-driven technique example known as Belief Rule Base shortly known as BRB, which is used when issue of uncertainty occurs [19–23]. As a result, the apps that we have included are extremely beneficial to disabled people.
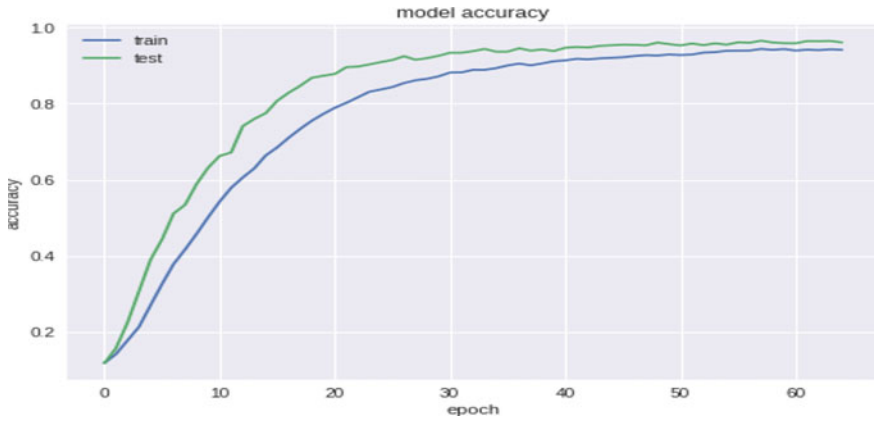
Fig. 9 Testing versus training accuracy of augmented model
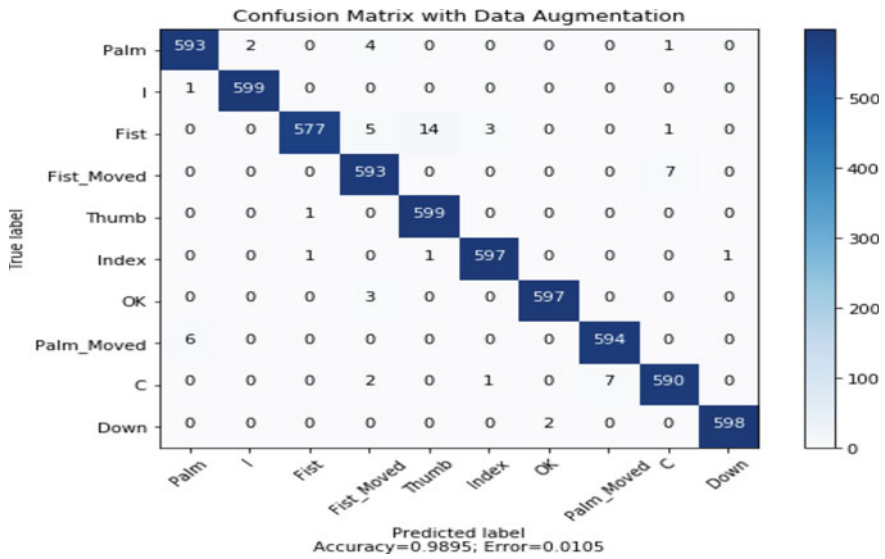
Fig. 10 Obtained confusion matrix of Hand Gesture Recognition Database

References

1. Wilson AD, Bobick AF (1995) Learning visual behavior for gesture analysis. In: Proceedings of international symposium on computer vision-ISCV, IEEE, pp 229–234
2. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436
3. Stergiopoulou E, Papamarkos N (2009) Hand gesture recognition using a neural network shape fitting technique. Eng Appl Artif Intell 22(8):1141–1158
4. Nguyen T-N, Huynh H-H, Meunier J (2013) Static hand gesture recognition using artificial neural network. J Image Graph 1(1):34–38
5. Chen Q, Georganas ND, Petriu EM (2008) Hand gesture recognition using haar-like features and a stochastic context-free grammar. IEEE Trans Instrum Meas 57(8):1562–1571
6. Molchanov P, Gupta S, Kim K, Kautz J (2015) Hand gesture recognition with 3D convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 1–7
7. Flores CJL, Cutipa AG, Enciso RL (2017) Application of convolutional neural networks for static hand gestures recognition under different invariant features. In: 2017 IEEE XXIV international conference on electronics, electrical engineering and computing (INTERCON), IEEE, pp 1–4
8. Zivkovic Z (2004) Improved adaptive Gaussian mixture model for background subtraction. In: null, IEEE, pp 28–31
9. Zivkovic Z, Van Der Heijden F (2006) Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recogn Lett 27(7):773–780
10. Grundland M, Dodgson NA (2007) Decolorize: fast, contrast enhancing, color to grayscale conversion. Pattern Recogn 40(11):2891–2896
11. Haralick RM, Sternberg SR, Zhuang X (1987) Image analysis using mathematical morphology. IEEE Trans Pattern Anal Mach Intell 4:532–550
12. Zhu Y, Huang C (2012) An improved median filtering algorithm for image noise reduction. Phys Proce 25:609–616
13. Manteco´n T, del Blanco CR, Jaureguizar F, Garc´ıa N (2016) Hand gesture recognition using infrared imagery provided by leap motion controller. In: International conference on advanced concepts for intelligent vision systems, Springer, 47–57
14. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhut-dinov R (2014) Dropout: a simple way to prevent neural networks from over-fitting. J Mach Learn Res 15(1):1929–1958
15. Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp 315–323
16. Dunne R, Campbell A (1997) On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation function. In: Process 8th Aust Conf on Neural Networks, vol 181, Melbourne, Citeseer, p 185
17. Bottou L (2010) Large-scale machine learning with stochastic gradient descent. In: Proceedings of COMPSTAT'2010, Springer, pp 177–186
18. Perez K, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621
19. Islam R, Hossain MS, Andersson K (2018) A novel anomaly detection algorithm for sensor data under uncertainty. Smooth Comput 22(5):1623–1639
20. Hossain MS, Rahaman S, Kor A-L, Andersson K, Pattinson C (2017) A belief rule based expert system for datacenter pue prediction under uncertainty. IEEE Trans Sustain Comput 2(2):140–153
21. Hossain MS, Ahmed F, Andersson K et al (2017) A belief rule based expert system to assess tuberculosis under uncertainty. J Med Syst 41(3):43

22. Hossain MS, Zander P-O, Kamal MS, Chowdhury L (2015) Belief- rule-based expert systems for evaluation of e-government: a case study. Expert Syst 32(5):563–577
23. Ul Islam R, Andersson K, Hossain MS (2015) A web based belief rule based expert system to predict flood. In: Proceedings of the 17th international conference on information integration and web-based applications and services, ACM, p 3

Community-Based Question Answering Site Using MVC Architecture for Rapid Web Application Development

D. V. S. S. Sujan, B. Lalitha, Ajay Reddy, A. Lakshmi Pathi, G. Sai Nikhil, and Y. Vijayalata

Abstract There are many social platforms that connect everyone across the globe. But still, we find a gap in connecting with our faculty, seniors, alumni, etc., who belong to the same college and community. It is important to know the perspective of others regarding any aspect related to the college. Making the students, faculty and alumni connected together will give a broader perspective of things that are relevant. We have developed a web-based application which allows students to ask questions and also give answers to the previously posted questions. And also giving special privileges to faculty and alumni to actively share their knowledge on a same platform enabling every person of the college to stay united and get updated with the current news. Any person belonging to the college can post questions and also answer the questions. There is no such application for any college in which everybody has the liberty to express their ideas, thoughts, and knowledge for other's problems. This idea not only makes us knowledgeable but also united.

Keywords Community-based question and answering forum · MVC · Keyword extraction

1 Introduction

Every individual across the globe has the eagerness to know something that they do not know. There is always something a person might not know about something in this world. Knowing something important puts you ahead of many people. In this era, it's all about getting updated through technology. There's always a lot to learn from the experienced people and taking insights from such people makes us stay at the top of the game.

D. V. S. S. Sujan (✉) · B. Lalitha · A. Reddy · A. L. Pathi · G. S. Nikhil · Y. Vijayalata
Department of Computer Science and Engineering, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, India
e-mail: dvsssujan@gmail.com

There are many platforms over the internet where people can share things, express their thoughts, share their knowledge and experience with anyone across the globe. These platforms enable an individual to gain insights and also will improve and change their lifestyle. But still, we find a gap in connecting with our faculty, seniors, alumni etc. who belong to the same college and community. It is important to know the perspective of others regarding any aspect related to the college. Making the students, faculty and alumni connected together will give a broader perspective of things that are relevant.

We would like to develop a web-based application which allows students to ask questions and also give answers to the previously posted questions. And also giving special privileges to faculty and alumni to actively share their knowledge on a same platform enabling every person of the college to stay united and get updated with the current news. Any person belonging to the college can post questions and also answer the questions.

The objective of our work is to:

- Classify queries based on keywords.
- Understand query patterns.
- Predict the answer for similar queries.

2   Literature Survey

There are many platforms over the internet where people can share things, express their thoughts, share their knowledge and experience with anyone across the globe. But still, we find a gap in connecting with our faculty, seniors, alumni etc. who belong to the same college and community. There are always some questions that keep running in our minds. And, such a question could be answered by someone who is a part of the same organization. There's always a possibility that multiple persons can have similar query running in their minds. So, answering a question could help multiple people at the same time.

So, there is a need of an application which allows users to gain and share knowledge from the community [1]. In the "Designing an MVC Model for Rapid Web Application Development", database access and routing components are built in the model. In this research paper, the built is done in PHP environment and can be further done any other environment [2]. In the "A Study on Q&A Services Between Community-based Question Answering and Collaborative Digital Reference in Two Languages", research is done on the community-based Q&A services and Collaborative digital reference where it is found that Community-based Q&A service sites provide more answers within shorter response times, and they are probably better suited to answer questions about everyday life [3]. In the "DOM-based Keyword Extraction from Web Pages", information and features including HTML and URL content are extracted from DOM tree of the page and then different scores are given to the text by their positions. And then from these top keywords are given out or given more importance.

MVC architecture that is implemented in the application is based on [1] paper. And, the idea to have a community-based Q&A service is based upon the [2] paper where it is found that community-based Q&A provide better insight than the global or digital application. The main objective of our paper is to tell that different user from the same community give out better information and would be more helpful than any other third-party applications. So, an application for the college community would be more beneficial and informative. And, the additional features which are implemented in the application, i.e., hate speech detection to remove any comments offending race, religion, inappropriate language, etc.; also, topics creation by the users and the notification system makes users life easier in accessing the followed content.

3 Methodology

The flow of the application is as follows:

1. Authentication

The application starts with creating login credentials for the user using which a user can login to our application. If the user did not login or create an account, the user can still view questions and answers but cannot add anything to it. If any user is creating a topic or question, they are recognized by their display picture.

2. Create Topics

The application contains some general topics initially. A user has the choice to create topics dynamically. While creating a topic, the user is required to enter description, title and also add a picture related to it. All the topics created are displayed in the topics section and the user can click on any topic to view the question, answers and description related to it.

3. Posting questions and answering them

A user can create questions under a particular topic and other users using the application can give answers to these questions. After entering a question or an answer, hate speech detection is done to identify any inappropriate or hate comments. If any found a warning pop out citing as a hint to change the input text.

4. Hate speech detection

We detect inappropriate text or foul language entered by the user, if found, a popup is raised to warn the user.
    The text which is inappropriate or hate is found through SVM.

5. Notification (Keyword Extraction)

We detect the keywords from the entered question or the answer or the topic from the user for which users receive notifications.

The application has a user interface which lets the user control the application through login/signup, create topic, question and answering. The user who is using the application is known by his display picture.

The architecture which is used for this application is MVC architecture (Model, View and Controller). Model view and controller architecture is used to develop web application user interfaces that separate the program logic into cluster of three interconnected elements. Firstly, when the application is accessed, it requires user credentials to access the site. The credentials are secured with 'bcrypt' algorithm with 10 rounds of hashing. Also, passport is used to support authentication for username and password. Once the user is logged in successfully, the user will be redirected to the topics page where they can create their own topic or access the topics. A topic is created as cards using bootstrap. In the homepage, the user will be getting the followed topics questions and answers.

A user can create his own questions and also can answer the questions which are already posted in the application. While answering or posting questions it undergoes hate speech detection which detects the hate present in it the flowchart of which is shown in Fig. 1. If found, a popup is raised giving a warning to the user. This process is done for the create topic section as well. And, whenever a question is posted all the users who are following the topic will get a notification and if an answer is given to a particular question, then a notification is sent to the user who posted the question.

Support vector machine

The support vector machine (SVM) is a supervised learning algorithm that can be used for both classification and regression problems. It is related to the principle of

Fig. 1 Flowchart of the methodology
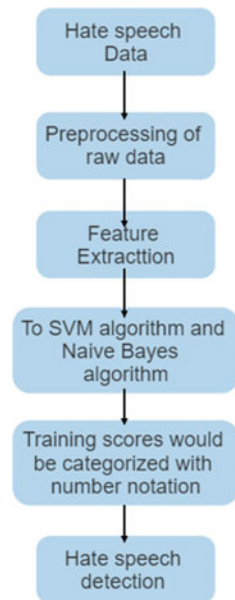
maximum margin. The SVM attempts to maximize the distance between the training data points, which it groups into two classes, while simultaneously attempting to minimize the distance from each data point to the hyperplane that separates them. The algorithm is optimized to find a hyperplane that separates the classes in the training set.

Keyword Extraction

The DOM-based keyword extraction technique is a way to extract keywords from the document object model. The DOM-based keyword extraction technique is not based on any particular algorithm, but instead relies on heuristics. The heuristics used by this method are based on the assumption that keywords are more likely to be contained in the text near the top of the document, and they are less likely to be found near the bottom of the document.

This assumption has been shown to be true for many documents, but not all documents contain keywords in this area. The DOM-based keyword extraction is a form of content analysis that extracts the keywords from a given text. This technique is considered to be more accurate than other forms of keyword extraction because it does not rely on the spelling or grammar of the sentence.

The DOM-based keyword extraction can be used to rank web pages and extract keywords for SEO purposes. Its architecture is shown in Fig. 2.
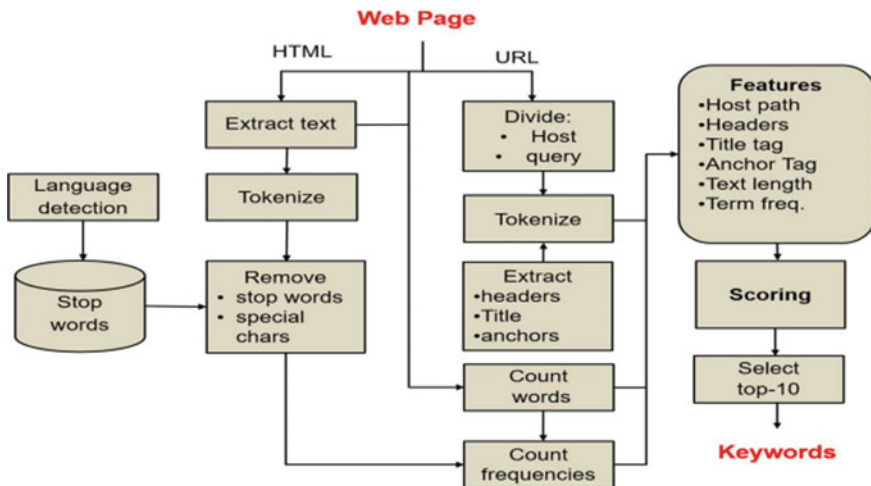
Fig. 2 DOM-based keyword extraction architecture

4 Existing Approaches

4.1 Yahoo Answers

Yahoo! Answers was a question-and-answer forum which is a community-driven application by Yahoo!. It provided a platform where users would ask questions and answers would be given by others. Answers submitted by others are voted based on the real content to increase their visibility.

4.2 Blurt It

Blurt it's far a Q&A Internet site in which human beings could ask questions and a community of normal users supplied might offer solutions to the one's questions based on their know-how or critiques. In this application, a user can publish query anonymously.

The application is full of Rampant racism, bullying and spam with no moderation.

4.3 Answers.Com

Answers.Com, previously called WikiAnswers, is an Internet-based expertise alternate. Answers.Com offers an Internet site. Where registered customers can interact with each other. There is not any way to verify that the answers are correct without doing back-up research.

5 Dataset

Figure 3 shows a sample data from the Hate speech dataset present in the Kaggle dataset which consists of almost 25,000 records and four attributes (count, hate_speech, offence, neither).

6 Experimental Results

After executing the application through terminal, the application is ready for execution and results. Firstly, a user needs to login to get into the application.

Once the user logs in (Fig. 4), the user will be redirected to topics section where topics can be created or existing topics can be accessed (Fig. 5).

Fig. 3 Dataset with HateSpeech

Fig. 4 Login page
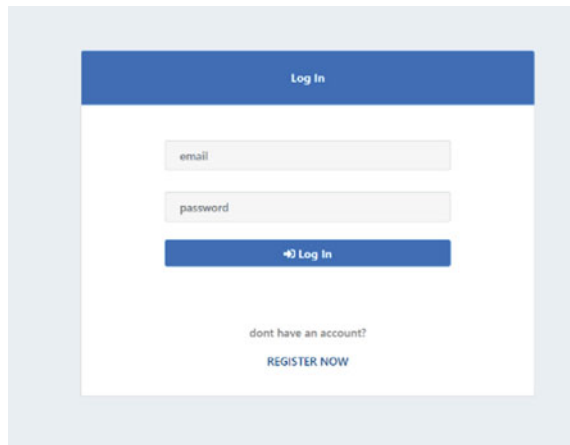
The user then can choose the topic and post question or answer to the existing question (Fig. 6). While entering questions or answers or creating topics, if any inappropriate or hate text found a pop up similar to Fig. 7 is raised.
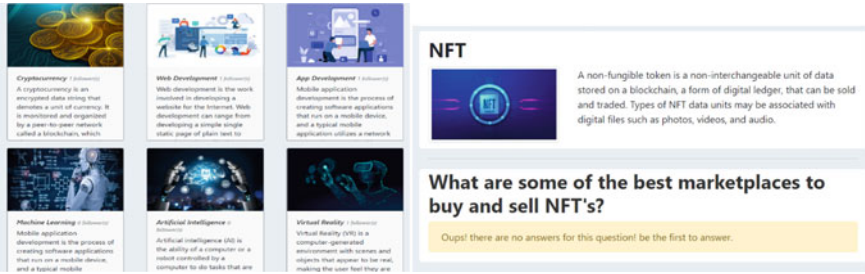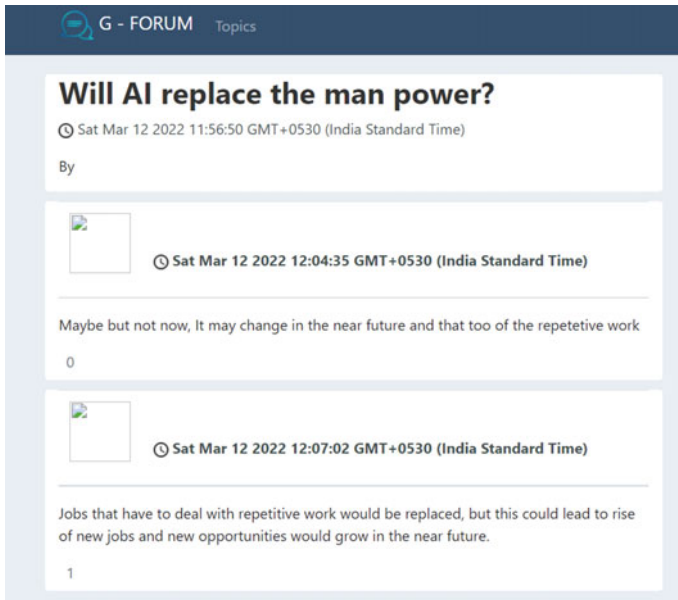
Fig. 5 Topics visible for end users

Fig. 6 Question and answer view page

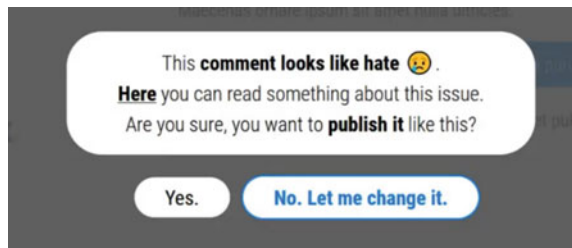**Fig. 7** Warning popup upon hate speech identification

Table 1 Results

TP rate	FP rate	Precision	Recall	*F*-measure	ROC	Area class
0.985	0.007	0.991	0.985	0.988	0.994	Not offensive
0.993	0.15	0.988	0.993	0.991	0.994	Offensive

Multiple users can answer to the same question and popular answer is shown at the top.

Model Information

The results have been tabulated in Table 1 and the model information is as follows:

Correctly Classified Instances—12,043   96.37%.
Correctly Classified Instances—12,043   96.37%.
Incorrectly Classified Instances—453   3.63%.
Kappa statistic   0.979.
K&B Relative Info Score 1,206,002.5861%.
K&B Information Score 11,983.9239 bits 0.9621 bits/instance.
Class complexity | order 0 12,439.7834 bits 0.9853 bits/instance.
Class complexity | scheme 22,197.4947 bits 1.7762 bits/instance.
Complexity improvement (Sf)—9861.375 bits—0.7891 bits/instance.
Mean absolute error 0.0149.
Root mean squared error 0.0971.
Relative absolute error 3.1054%
Root relative squared error 18.961%
Total Instances 12,497.
Accuracy of model 98.27%

7 Conclusion and Future Scope

There's always a need for answering questions wandering in our minds. We all need a place to express ourselves, share our experience with those who need it. There's always a scope to react and interact with others over a platform. Any experienced person could help a newbie. Our project will help the members of the college to indirectly interact with others by posting questions and getting answers from faculty and alumni. Edge-cutting technologies are being used to develop this application so that no user will have any problem of compatibility across any device.

G-FORUM is a place or a platform which connects students, faculty, alumni, and other members of the college at one place.

The future scope of the project is definitely adding more value to the project so users can use the platform more effectively. There are some future scopes that follows:

- Adding more machine learning algorithms, improving user experience.
- Giving special privileges to different users of the application.
- Commercializing the project to other colleges by adding a network.

Exploring dynamic hate speech detection embedded in websites by not restricting to a limited feature space.

References

1. Wu D, He D (2013) A study on Q&A services between community-based question answering and collaborative digital reference in two languages. In: iConference, Fort Worth, TX, USA
2. Rezaei M, Shah H (2019) DOM-based keyword extraction from web pages. In: Artificial intelligence, information processing and cloud computing (AIIPCC) 2019
3. Priyambiga R, Sonkamble RG, Machhale GG, Mulla R (2016) Modern question answering forum. In: Techno societal 2016 international conference on advanced technologies for societal applications
4. Mathew B, Saha P, Yimam SM, Biemann C, Goyal P, Mukherjee A (2021) HateXplain: a benchmark dataset for explainable hate speech detection. In: AAAI (Association for the advancement of artificial intelligence)
5. Altar A, Samuel, Pop DP (2014) Designing an MVC model for rapid web application development

Automatic Alert and Triggering System to Detect Persons' Fall Off the Wheelchair

Syed Musthak Ahmed, Sai Rushitha, Shruthi, Santhosh Kumar, Srinath, and Vinit Kumar Gunjan

Abstract Sensor-based human activities need a lot of attention in the driving technologies of the Internet of Things (IoT) because it should set the human activities (make and perform personal assisting tasks) that should be performed well. Taking care of physically disabled persons and preventing them from falling from a wheelchair are vital tasks to be performed. Scrutinizing these kinds of accidents via surveillance systems based on CCTV cameras would not prevent them from falling from the wheelchair. In the present work, a solution to detect and signal the situation, an intelligent and cost-effective fall detection system is presented. This module utilizes the benefits of the advanced technology, i.e., the Internet of Things (IoT). Here, an ultrasonic sensor is fixed on to the wheelchair to detect the obstacles such as pits, stairs, and an accelerometer to detect human movements. These sensors are connected through a microcontroller to transmit acceleration data continuously. The system monitors the falls and abrupt changes in the person's movement. A sudden jerky change in the system is treated as a crash or fall. The system senses these changes and automatically triggers a warning via the Wi-Fi connection and sends information to the near ones about the situation via the Blynk application.

Keywords Fall detection · Accelerometer · Blynk application

1 Introduction

In the world at large, 33% of senior citizens are tumbling down in their resident places yearly with respect to their illness and health conditions. Taking care of their well-being is a necessary action to be made. Their health conditions are the main reasons, for falling significantly [1]. The wheelchair is controlled by a touch screen-based navigation system along with a GPRS system for location determination and

S. M. Ahmed (✉) · S. Rushitha · Shruthi · S. Kumar · Srinath
Department of ECE, SR Engineering College, Warangal, Telangana, India
e-mail: syedmusthak_gce@rediffmail.com

V. K. Gunjan
CMRIT, Hyderabad, Telangana, India

GSM to communicate in abnormal events like falling [2]. With the advancements in the technologies like IoT, it is possible to develop elderly caring smart devices with features such as forward, backward, upstanding, and so on [3, 4]. For independent living, there is a need for the mobility of aged adults and physically challenged people to maintain their health and to maintain their cognitive abilities. The impairments are of different types which are visual, auditory, mobility, and cognitive [5–8]. A prototype of a smart wheelchair for people of different disabilities is developed to detect fall is presented in [9–11]. Hand gestures controlled movement of the chair is also presented in their work. They incorporated an accelerometer, voice-control using ultrasonic sensors, and a GSM module in their implementation. Automatic wheelchair to navigate without any human assistance using technologies like MEMS, GSM, and Wi-Fi to support the disabled is presented in [12–16]. A four-wheeled omnidirectional wheelchair interfaced with a myoelectric controller, driven by a holonomic drive system, providing greater flexibility compared to conventional powered wheelchairs are presented in [17–19]. A semi-autonomous navigation of wheelchair mounted with sonar's and cameras to extract sensor information for impaired peoples' mobilization is presented in [20].

2 Existing Methods

There are several works reported on auto navigation of wheelchairs to assist the disabled [13–20]. Previous works also included physio-operations, cognitive development, and massage operation [5, 6]. Ahmed et al. in their work incorporated MEMS technology for navigation purposes. This is shown in Fig. 1a–c They included a wireless system that helps to navigate the physically disabled and aged adults [5]. The movement of the wheelchair is supported by means of a DC geared motor in response to the accelerometer. Physiotherapy operation for the neck and back was incorporated. Physiotherapy operations were performed by means of an actuator, a rope, and a pulley [6]. Actuators operation was used to balance the distribution of load between the hands and the legs. Vibration motors perform massage operations for the legs, back, and neck. A potentiometer is used to control the amount of vibration. The Raspberry Pi CPU is used to carry out cognitive functions and is monitored on a display using Wi-Fi. The ability to recognize and respond to pictures, numerals etc. determine a person's cognitive level. The ultimate goal is to automate as much as possible to reduce daily work at the same time taking care of the elderly at home. In [13], infrared radiation (IR), joysticks, eye tracking, and a speech restructuring system are included in the control wheelchair. Low-intensity infrared photons were attempted to transit on the eyes. Depending on the movement of the eyelids, the voltage level in the IR receiver varied that helped in wheelchair navigation.
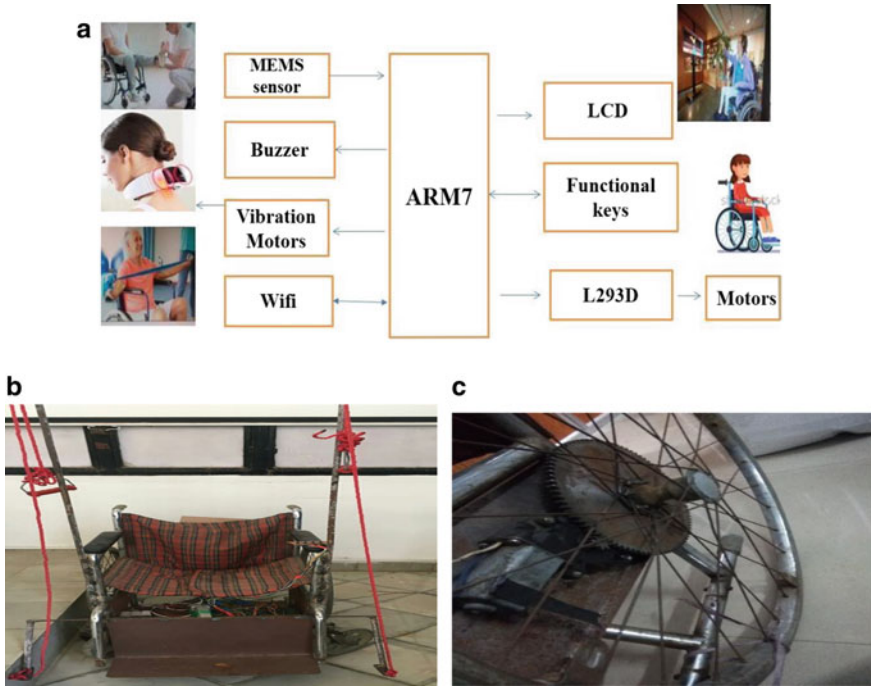
Fig. 1 **a** Block diagram of auto navigation wheelchair **b** wheelchair for auto navigation with physiotherapy [5, 6] **c** wheelchair with motor and gear system [5, 6]

3 Proposed Method

In the proposed work (Fig. 2), a technique to screen our old ones for their well-being and security in work carried out. Because of shortcomings and weak joints, they have an incredible danger of tumbling down. For this reason, we propose a smart fall identification framework. The framework utilizes an accelerometer to identify individual development, and it is mounted on a wheelchair for identification. The sensor is associated with a microcontroller to communicate the information continually. An unexpected sudden change with a jerk in the framework is treated as a fall. The framework identifies an individual who has fallen and consequently triggers an alert through wireless associated with alerting the neighbors and/or family of the individual with regards to the circumstance in a split second. Figure 3a shows the block diagram of developed model while Fig. 3b the practical implementation.
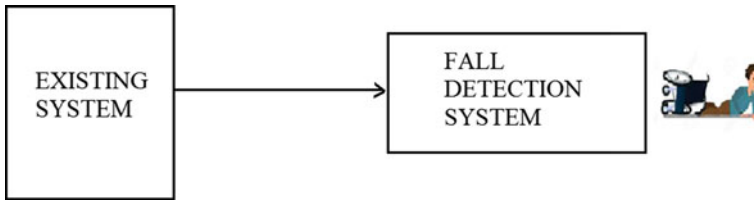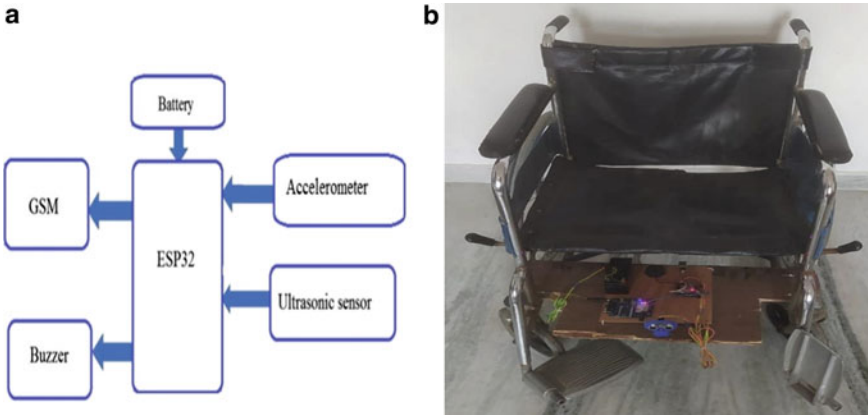
Fig. 2 Development over the existing system

Fig. 3 a Block diagram of the proposed method **b** wheelchair with fall detection

4 Objectives

The purpose of the developed product is to

1. Assemble a low cost, fall detection smart wheelchair system.
2. Trace the presence of an obstacle in the path of navigation.
3. Prevent fall of wheelchair and protecting the person from falls.

5 Methodology

The following steps are followed in developing the model

1. A wheelchair with auto navigation circuit is implemented
2. An ultrasonic sensor is fixed to detect the obstacle and a buzzer to alert of fall
3. An accelerometer is fixed to detect movements and to signal the instance of fall
4. A Wi-Fi along with GSM module, and Blynk application to inform the care taker.

6 Results and Discussion

The wheelchair to detect the fall due to any obstacle is being developed. The work is studied under four directions of fall, namely left, right, front, and back. These are explained by a pictorial view in Fig. 4 along with alert messages.

1. Left Fall Detected: Here, the framework identifies whether the person has fallen or not; if the person falls left side, an alert message as "Alert!! Left Fall Detected" is sent to the caretakers by using the GSM module and Blynk application.
2. Right Fall Detected: Here, the framework identifies whether the person has fallen or not; if the person falls right side, an alert message as "Alert!! Right Fall Detected" is sent to the caretakers by using the GSM module and Blynk application.
3. Forward Fall Detected: Here, the framework identifies whether the person has fallen or not; if the person falls forward, an alert message as "Alert!! Forward Fall Detected" is sent to the caretakers by using the GSM module and Blynk application.
4. Backward Fall Detected: Here, the framework identifies whether the person has fallen or not; if the person falls backward, an alert message as "Alert!! Backward
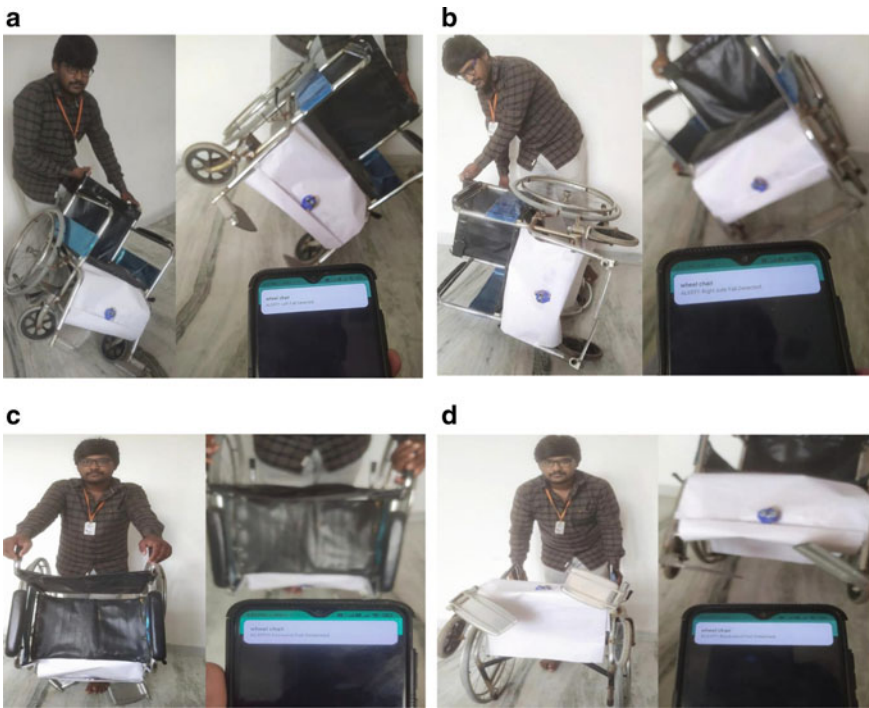

Fig. 4 **a** Left fall detected **b** right fall detected **c** forward fall detected **d** backward fall detected

Fall Detected" is sent to the caretakers by using the GSM module and Blynk application.

Thus, by adding this feature on to the wheelchair to the existing model the wheelchair is made smarter in its navigation, care and also in terms of safety.

7　Conclusion

Many falling detection systems have experimented earlier with sophisticated and some other specific sensors. The poor design part, cost-effectiveness, and complexity in usage are the major drawbacks that customers could not attain. Accelerometer helps to detect a person falling from a wheelchair and sends a notification to caretakers with the help of the GSM module. The model functions admirably with its obstruction location capacities accomplished by utilizing ultrasonic sensors. An alarm will quickly alert the people who are nearer, and it helps them to react quickly according to the circumstances. This system helps not only senior citizens but also disabled people, those who crawl, those who walk with the help of a stick, those who have severe and chronic joint/tissue problems and also with stiffness in joint movement, or who have no mobility, voluntary movement, or tremor or have weak bones.

References

1. ArkaprabhaLodh DG, Ghosh D (2018) Accelerometer and arduino based gesture controlled robocar. Int J Innovative Res Sci Eng Technol 7(8):2347–6710
2. Siddharth PD, Deshpande S (2016) Embedded system design for real-time interaction with Smart Wheelchair. In: IEEE symposium on colossal data analysis and networking (CDAN), pp 1–4
3. Gia TN, Tcarenko I, Sarker VK, Rahmani AM, Westerlund T, Liljeberg P, Tenhunen H (2016) IoT-based fall detection system with energy efficient sensor nodes. IEEE Nordic circuits and systems conference (NORCAS), pp 1–6
4. Tacconi C, Mellone S, Chiari L (2011) Smartphone-based applications for investigating falls and mobility. In: 5th International conference on pervasive computing technologies for healthcare (Pervasive Health) and workshops, pp 258–261
5. Ahmed SM, Yasmeen A, Jagadeesh Babu B (2018) Wheelchair with auto navigation for adults with physio and cognitive impairments. Int J Innovation Technol Exploring Eng (IJITEE) 8(252):458–462. ISSN: 2278-307
6. Ahmed SM, Shireen A, Jagadeesh Babu B, Shruti (2020) Powered wheelchair for mobility with features to address physical strength, cognitive response and motor action development issues. Published in lecture notes in electrical engineering LNEE, vol 601, Springer book series, pp 1110–1117
7. Mauldin TR, Canby ME, Metsis V, Ngu AHH, Rivera CC (2018) SmartFall: a smartwatch-based fall detection system using deep learning. Sensors 2018 18(10)
8. Waheed SA, Khader PSA (2017) A novel approach for smart and cost effective IoT based elderly fall detection system using Pi camera. In: IEEE International conference on computational intelligence and computing research (ICCIC), pp 1–4

9. Vora H, Gupta A, Pamnani C, Jaiswal T (2020) Multimodal smart wheelchair integrated with safety alert system. Int J Eng Adv Technol (IJEAT) 9(4)
10. Sposaro F, Tyson G (2009) iFall: an android application for fall monitoring and response. In: Annual international conference of the IEEE engineering in medicine and biology society, pp 6119–6122
11. Lavanya KN, Shree DR, Nischitha BR, Asha T, Gururaj C (2017) Gesture controlled robot. In: IEEE International conference on electrical, electronics, communication, computer, and optimization techniques (ICEECCOT), pp 465–469
12. Madane Mugdha R, Agarwal RV, Ghare Radha P, Thorat PS (2015) Gesture control wireless wheelchair prototype. Int J Eng Res Technol (IJERT) 04(04)
13. Megalingam RK, Chacko C, Kumar BP, Jacob AG, Gautham P (2016) Gesture controlled wheel chair using IR-LED TSOP pairs along with collision avoidance. In: International conference on robotics and automation for humanitarian applications (RAHA), pp 1–7
14. Simpson RC, LoPresti EF, Cooper RA (2008) How many people would benefit from a smart wheelchair? J Rehabil Res Dev 45:53–71
15. Simpson RC (2005) Smart wheelchairs: a literature review. J Rehabil Res Dev 42(4):423–436
16. Richard CS, Levine SP (2002) Voice control of a powered wheelchair. IEEE Trans Neural Syst Rehabil Eng 10(2)
17. Madarasz RL, Heiny LC, Cromp RF, Mazur NM (1986) The design of an autonomous vehicle for the disabled. IEEE J Robot Autom 2(3):117–126
18. Ding D, Cooper RA (2006) Electric powered wheelchairs. IEEE Control Syst Mag 22–34
19. Bourhis G, Moumen K, Pino P, Rohmer S, Pruski A (1993) Assisted navigation for a powered wheelchair. Systems engineering in the service of humans: proceedings of the IEEE international conference on systems man and cybernetics 3:553–558
20. Argyros A, Georgiadis P, Trahanias P, Tsakiris D (2002) Semi-autonomous navigation of a robotic wheelchair. J Intell Robot Syst 315–329

Dietary Assessment by Food Image Logging Based on Food Calorie Estimation Implemented Using Deep Learning

Syed Musthak Ahmed, Dayaala Joshitha, Alla Swathika, Sri Chandana, Sahhas, and Vinit Kumar Gunjan

Abstract According to recent studies across the world, it is observed that a healthy diet is a key to have sound health and body. Nowaday's people are more concerned with their diet than ever before. Obesity has become a significant health problem across the world because it is associated with many of the leading causes of ill health such as chronic diseases including diabetes, heart stroke, renal disease, and cancer. The most effective way to prevent obesity is through food intake control which involves understanding food ingestion like the calories in each and every intake. In our daily diet, fruits, and vegetables play a vital role because they are the major source of energy. A nutritious diet present in fruits and vegetables can help us to tackle different health problems like cancer, diabetes, heart diseases, etc. With the advancements of science and technology, it's now viable to construct a singular identification system for keeping track of day-to-day diet with fruit and vegetable calorie intake that can be very useful to maintain good health without any expert dietitian advice. This paper proposes a deep learning model which ensure both user flexibility and provide high accuracy in classifying and predicting calories present in fruits and vegetables. In order to quantify the intake of food and make one fit with good health and longevity we are suggesting dietary plans for different diseases like high blood pressure, rheumatoid arthritis, high cholesterol, etc.

Keywords Convolutional neural networks · MobilenNetV2 · Deep learning · Image recognition · Data augmentation · Web scraping

---

S. M. Ahmed (✉) · D. Joshitha · A. Swathika · S. Chandana · Sahhas
Department of ECE, SR Engineering College, Warangal, Telangana, India
e-mail: syedmusthak_gce@rediffmail.com

V. K. Gunjan
CMRIT, Hyderabad, Telangana, India

1   Introduction

Health is one of the foremost important aspects of a person's life. Obesity has become a significant health concern in many parts of the world. High calorie food intake is often harmful and leads to obesity. As per medical research, 1.9 billion people aged above 18 years are suffering from being overweight including children are more likely to consume nutrition and other junk foods than before. It leads to health problems like heart diseases, diabetes, blood pressure, and cancer. To overcome these problems people are paying more attention on food consumption and a diet chart or plan are planned to maintain their weight is followed by every individual. In each and every diet plan fruits and vegetables contribute to an essential part because major sources of energy, vitamins, fiber, plant chemicals, and nutrients. A diet with fruit and vegetables has high nutrients that can help one to prevent from cancer, diabetes, heart diseases, etc.

2   Literature Review

Researches in the literature have often focused on different aspects of the food recognition problem. Many works address the challenges in the recognition of food by developing recognition strategies that differ in terms of features and classification methodologies. In [1] a system is proposed for the food recognition and calorie estimation using single shot Multibox detector (SSD) algorithm. They used object detection to estimate calorie count of some famous Chinese dishes along with Western dishes. Jasmine Minija and Sam Emmanuel [2] proposed a model which predicts calorie content present in a particular food item. In their work, they incorporated CSW-WLIFC algorithm for food segmentation. Pooja et al. [3] proposed a model where in the input image were passed through the MathWorks Image processing Toolbox to extract features. These compressed images were passed through the classifier to predict food item and regressor to estimate size and calorie content. Kuhad et al. [4] proposed a single and multiple mixed food object recognition using deep learning and SVM. Finger-based calorie measurement, distance measurement methods were applied for calorie estimation. Jasmine Minija and Sam Emmanuel [5], proposed CSW-WLIFC based segmentation process using CSW-based kernel function for feature extraction and WLM-NN classifier for classification for segmented food items and they improved performance with values of 0.99 for segmentation accuracy. In [6], Özsert Yiğit and Özyildirim, proposed a method using deep convolution neural network (DCNN) structure with Stochastic gradient descent, Nesterov's accelerated gradient, and adaptive moment for estimation. Liang et al. [7] proposed the food calories count for weight prediction model using object detection technique to identify the food and estimate calorie based on the images. The results of calorimeter R2, RMSE were about 0.95 and 43 while MSE was about 32, and the draw error rate is about 9%. Auto food log record system which automatically estimate

calories of food using machine learning methods Bayesian; support vector machine is presented in [8]. In [2], food processing and recognition techniques are presented using *K*-Nearest Neighbor (KNN) algorithm which is a very simple, understandable, and highly efficient method in determining the calorie estimation. Ege and Yanai [9] proposed a system to estimate food calorie from a captured food image by simultaneous learning of categories, ingredients and cooking directions using multi-task CNN. In [5], Bayesian Fuzzy Clustering (BFC) and IpCA-DBN was used for recognition and feature extraction. In [10], a deep learning-based visual food recognition algorithm is designed to employ edge computing-based service computing paradigm to overcome few inherent problems of traditional mobile cloud computing paradigm, such as unacceptable system latency and low battery life of mobile devices. Dense SIFT method and extracted visual vector using *K*-means clustering technique is being adopted in [11]. To support their work, support vector machine classifier was used to classify the food image and measures the carbohydrate level. Calorie measuring system using special calibration card technique to measure calories and nutrition in every day meals to help patients and dieticians by measuring and manage daily food intake is presented in [12]. In [13], artificial neural networks (ANNs) for calorie estimation to various fruits on a plate using random forest algorithm along with structural features is presented.

3 Proposed System

In the proposed work to obtain better results a deep learning model, namely, convolution neural network (CNN) is used to improve the accuracy. In present work, a web application is designed to provide users with rapid and accurate results about the number of calories present in the image along with the fat content present in the fruits and vegetables that is consumed by an individual in their diet. Once, the image is uploaded and submitted in JPEG or PNG format, the uploaded image will first undergo pre-processing by resizing the image based on the model. After preprocessing, the image is optimized and sent to the next stage called classification, where CNN classification is done. Here, extraction of features like shape, color, texture, etc., are done. Once the classification is done, the calorie and fat count of predicted fruit/vegetable are obtained. On the other hand, along with the calorie count and fat content the user can check weekly diet plans for different diseases. The complete process is done within few milliseconds thereby provided fast answers with optimized accuracy. The various steps adopted in our model is explained in the following sections.
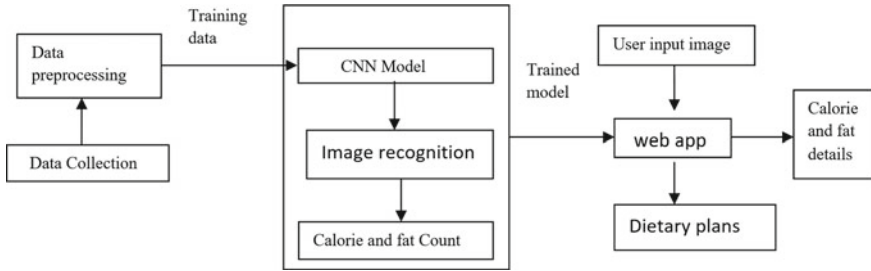
Fig. 1 System architecture of proposed model

3.1 System Architecture

The system architecture of the implemented is shown in the Fig. 1. Here, we the calorie content present in fruits and vegetables is estimated by means of images collected from the dataset. Based on the type of disease a diet plan is recommend. In present work, the diet plan is created for few diseases like blood pressure, cholesterol, diabetes, weight loss, PCOD, thyroid, and rheumatoid arthritis.

3.2 Data Collection

In implementation of our project, "Fruits and vegetable Image recognition" dataset is used. In the present dataset, 36 different categories are taken which include fruits like banana, apple, pear, orange, grapes, etc., and vegetables like cucumber, carrot, potato, etc., as shown in Fig. 2.

3.3 Data Pre-processing

In order for the recognition to take place an RGB file from a directory is loaded into our system. The image is then passed and resized to 224×224 dimensions to speed up the performance of model. In order to increase the size of the sample data taken data augmentation technique is used. The dataset consists of 36 different classes and each class is divided into training, testing, and validation images. Altogether, there are 3825 samples.
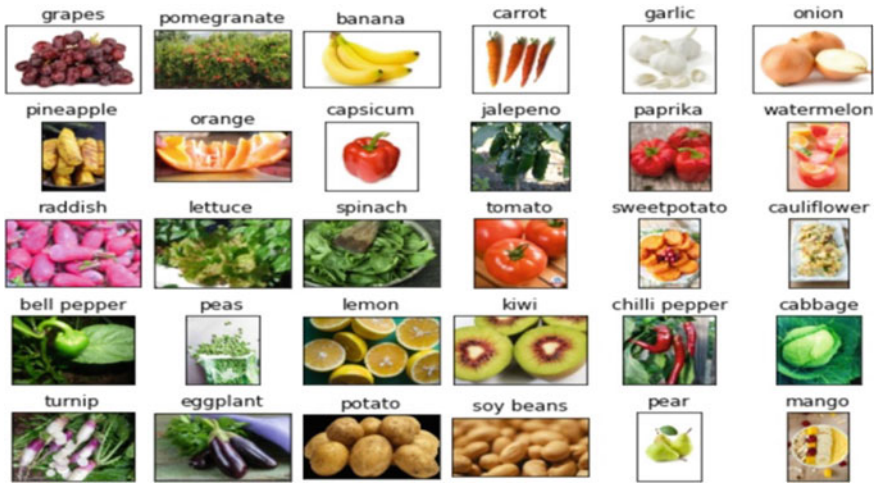
Fig. 2 Sample dataset images

3.4 Feature Extraction and Classification

In the present system, Fig. 3 MobileNetV2 architecture is incorporated. It is a lightweight deep neural network with less complexity and low computation costs. Thus, it's likely used for mobile and embedded applications. MobileNetV2 consists of four blocks. One is a residual block with a stride of 1, the second block with a stride of 2 for downsizing the image. There are three different layers for both types of blocks. The first layer is 1×1 convolution with a RELU6 activation function, the second layer is depth wise convolution and the third layer is another 1×1 convolution but without any nonlinearity. It is claimed that if RELU is used again, the depth networks only have the power of a linear classifier on the non-zero volume part of the output domain.
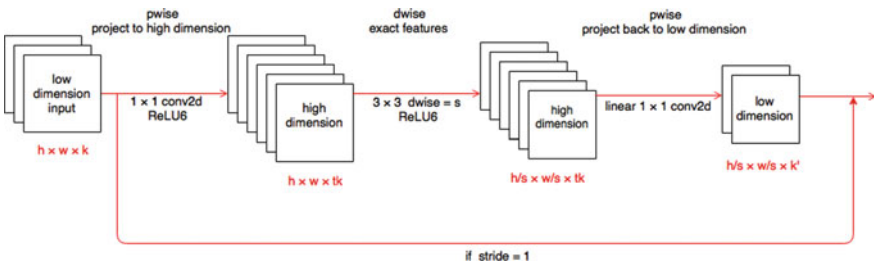

Fig. 3 MobileNetV2 architecture

Fig. 4 Calorie and fat content for carrot

3.5 Calorie Extraction

Once, the image of fruit/vegetable is analyzed and detected, our classifier is used to estimate the calorie content of the classified fruit/vegetable through web scraping. It also gives the fat content along with calories.

4 Results and Discussion

The web application developed provides the calorie count and the information regarding various clinical diseases, their symptoms, causes, complications and treatment methods along with the weekly dietary plan. The frontend-backend of web is handled by Streamlit framework. The developed model captures the input image and categorizes it into fruit/vegetable, classifies it and estimates the calorie and fat content present in the item. Figures 4 and 5 shows the categorization process carried out by our model. Figure 4 signifies the vegetable classification process while the Fig. 5 for fruit classification.

Table 1 describes the information regarding various clinical diseases, their symptoms, causes, complications, and treatment methods along with the weekly dietary plan.

5 Conclusion and Future Scope

In the present work, a model to assist one to lead a healthy and longevity by resolving dietary issues is presented. Here, a model is developed by identifying clinical problem and provide a suitable diet plan. The website developed provides necessary food

Fig. 5 Calorie and fat content for banana

Table 1 Information regarding diet plans

Clinical terminology (main page)	Type of diseases	Sub pages	Description
Cardio (Cardio.html)	High blood pressure Low blood pressure Cholesterol	Diethbp.html Dietlbp.html DietCl.html	Weekly diet plans and information on HBP Weekly diet plans and information on LBP Weekly diet plans and information on cholesterol
Obesity (weight loss)	1200 cal diet 1600 cal diet 1800 cal diet 2500 cal diet	Diet1200.html Diet1600.html Diet1800.html Diet2500.html	Weekly diet plans on 1200 cal diet Weekly diet plans on 1600 cal diet Weekly diet plans on 1800 cal diet Weekly diet plans on 2500 cal diet
Diabetes (Diabetes.html)	Type1 Diabetes Type2 Diabetes Gestational diabetes	Diettype1.html Diettype2.html Dietges.html	Weekly diet plans and information on Type 1 Weekly diet plans and information on Type 2 Weekly diet plans and information on gestational
Rheumatoid arthritis (Dietindex.html)	Rheumatoid arthritis	dietRA.html	Weekly diet plans and information on RA
Thyroid (Dietindex.html)	Thyroid	dietthyroid.html	Weekly diet plans and information on Thyroid
PCOD (Diabetes.html)	PCOD	dietpcod.html	Weekly diet plans and information on PCOD

information to diseased by providing data of the amount of calorie and the fat content present in the food item based on images. The identification of food item is carried out using CNN algorithm as it provides highest accuracy. The developed model is unique with a novel idea to prevent people from major health diseases such as diabetics, blood pressure, and many more chronic diseases. In the present work, calorie and fat estimation is carried out by taking few samples of fruits and vegetables. However, the work can be further extended by adding all varieties of fruits and vegetables and estimating their calorie and fat contents. Further, non-vegetarian foods may also be categorized and provided with necessary protein contents as this is also one of necessary nutrient needed for a diet plan.

References

1. Hu H, Zhang Z (2020) Image based food calories estimation using various models of machine learning
2. Jasmine Minija S, Sam Emmanuel WR (2017) Neural network classifier and multiple hypothesis image segmentation for dietary assessment using calorie calculator, Taylor and Francis
3. Pooja H et al. (2016) Food recognition and calorie extraction using bag-of-SURF and spatial pyramid matching methods. Int J Comput Sci Mob Comput 5(5):387–393
4. Kuhad P (2015) A deep learning and auto-calibration approach for food recognition and calorie estimation in mobile e-health, School of EE and CS
5. Minija Jasmine S, Sam Emmanuel WR (2019) Imperialist competitive algorithm—based deep belief network for food recognition and calorie estimation, Springer-Verlag GmbH Germany
6. Özsert Yiğit G, Özyildirim BM (2018) Comparison of convolutional neural network models for food image classification, Taylor and Francis
7. Liang H, Gao Y, Sun Y, Sun X (2018) CEP—calories estimation from food photos, Taylor and Francis
8. Sasano S, Han XH, Chen YW (2016) Food recognition by combined bags of color features and texture features. In: 2016 IEEE image and signal processing, biomedical engineering and informatics (cisp-bmei), international congress on 2016
9. Ege T, Yanai K (2017) Image-based food calorie estimation using knowledge on food categories, ingredients and cooking directions. In: Proc. thematic workshops
10. Liu C, Cao Y, Chen G, Vokkarane V et al. (2017) A new deep learning recognition system for dietary assessment on an edge computing service infrastructure. IEEE Trans Serv Comput 99
11. Velvizhy P, Pavithra, Kannan A (2014) Automatic food recognition system for diabetic patients. In: 2014 IEEE advanced computing (ICoAC), 2014 sixth international conference on advanced computing (ICoAC).
12. Kavita S, Pavithra S (2015) Performance analysis of nutritional contents in food images using SARAN. ARPN J Eng Appl Sci
13. Alghazo JM, Latif G, Alzubaidi L, Elhassan A (2019) Multi-language handwritten digits recognition based on novel structural features. J Imaging Sci Technol 63(2):20502–20511

Neck Gesticulate Based Vehicle Direction Movement Control System to Assist Geriatrics

Syed Musthak Ahmed and A. Alekhya

Abstract Transportation is an essential need for every individual to reach their destination. People who are physically fit are really fortunate as they can make their own drive on the vehicle while physically challenged and/or partially disabled people cannot drive a vehicle with the assistance of steering. As a result, a suitable solution for such a population to make a livelihood is required. The quadriplegic patients move independently from one place to another place with the help of the neckband by tilt movement of their head. The person driving the car will be supplied with a gadget that is worn around their neck and allows them to move the steering forward and backward without exerting any mental or physical effort. The prototype vehicle is driven by two 60 RPM geared motors, and it can make fast turns to the left, right, front, and back. Here, ATmega328p MCU is used as controller and four switches with ON/OFF control in the prototype model. Depending on the person neck movement the switch is press ON which makes the vehicle to move either a forward or backward and left or right directions.

Keywords Arduino Uno · Microcontroller · DC Motor · RF transmitter · RF receiver

1 Introduction

One-fifth of the world's population is projected to have serious disabilities, such as vision, hearing, hands, and legs. Limb disability is a type of impairment that can be caused by a variety of factors, including congenital deformity, war, and diseases like diabetes. Sportspeople's lower limbs take a lot of punishment while they play, and they're always at risk of serious injuries. There are a vast number of persons in the globe who have serious physical impairments, as well as the elderly, who have substantial trouble completing even fundamental functions like movement, speaking, and writing. Quadriplegics, who have been paralyzed across a major portion of their

S. M. Ahmed (✉) · A. Alekhya
Department of Electronics and Communication Engineering, SR University, Warangal, Telangana, India
e-mail: musthak.ahmed@sru.edu.in

body, are the most severely impacted group of physically handicapped. Any task requiring even a tiny amount of power is incredibly difficult for these folks. We propose a gesture-based wheelchair to assist physically challenged and elderly people in navigating inside their homes with minimal effort [1].

Voice instructions and the movement of hand are used to drive a wheelchair using a system based on speech and gestures. The system's two main components are the voice recognition module and the MCU. The most common method for identifying speech commands in the voice detection module is hidden Markov models. The MEMS sensor detects the angle of the hand and sends voltages to the microcontroller in accordance with it [2, 3].

Sundara Siva Kumar et al. [4] used MEMS technology to produce a wheel chair control that is beneficial to physically impaired people using their hand movement or hand gesture recognition. One of the most significant breakthroughs toward the inclusion of severely physically challenged individuals is the employment of a motorized wheel chair with advanced navigational intelligence. For those with arm or hand disabilities, driving a wheel chair in a home setting is problematic. The wheel chair was designed to address the aforementioned issue by permitting end users to perform only appropriate movements and complete certain important daily tasks.

According to different studies and surveys, having a method of autonomous movement improves both children and adults. Despite the fact that many handicapped people are satisfied with conventional manual or motorized wheelchairs, a segment of the disabled community finds it challenging or unable to use wheelchairs on their own. Many researchers have experimented with a variety of technologies in order to make wheelchairs more accessible to this demographic [5].

Significantly for a normal individual, driving a wheelchair in a home setting is tough, and it gets even more challenging for those who have lost use of their arms or hands. One of the most significant breakthroughs toward integrating severely physically challenged and mentally handicapped persons is the use of motorized wheelchairs with high navigational intelligence [6]. The motorized wheel chair is a remarkable technical breakthrough that has the potential to improve the lives of persons who are physically or intellectually challenged.

An electronic wheelchair cannot yet suit the demands of patients such as paraplegics, who actually spend the majority of the time in bed. This study describes the control system of a voice-activated wheelchair that may convert into a sitting, lying down, or standing posture. The wheelchair's hardware circuit and software program for speech control have been verified and tested for its functionality, and the recognition accuracy for voice activation for the same individual are satisfactory [3].

Ahmed et al. [7, 8] in their works developed an electrically powered customizable, low cost auto navigation wheel chair using MEMS technology. In his two different works he added navigation along with physic and cognitive developments to assist the aged people and also providing working group to monitor their caring ones from work place and keeping a watch on emergency needs. In [9–15], works on navigation using wheel chair are being discussed. The ultimate aim being to help the disabled with less arm power/voice control etc.
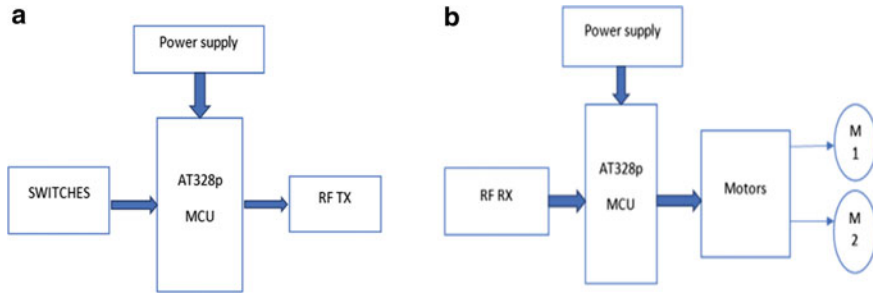
Fig. 1 **a** Transmitter **b** receiver

2 Block Diagram of Prototype

In both the transmitter and receiver sections of the planned system, there is a microcontroller. This vehicle moves in a certain direction by using a neck movement as an input signal. The switches are used to track the motions controlled by the switches are fitted to the neckband. The microcontroller receives the fluctuations of those signals as input, and the microcontroller is programmed to make judgments based on those variations, which then govern the vehicle's movement If the person tilts his head in right or left directions vehicle will move in the right or left directions and vice versa.

The block diagram of a transmitter is shows in Fig. 1a, it is built using Arduino Uno, switches and RF transmitter, while Fig. 1b shows the block diagram of a receiver, built using Arduino Uno, RF receivers and L293 motor control unit. RF receiver will receive the data and give this input to the microcontroller and the controller will regulate the directions of the neck movement. This project uses 7805 three-terminal voltage regulator.

3 Hardware and Software Requirements

3.1 Proteus Software

Proteus is a program that exclusively takes hex files. Once the machine code has been translated to hex code, this software dump the hex code into a microcontroller. Conversion of the given assembly language/High level language program into machine code is done by Keil Compiler.

3.2 Arduino Uno

Arduino Uno is a open source Microchip ATmega328p microcontroller. The Arduino Uno board is shown in Fig. 2, this board has the following features-DAC and ADC

Fig. 2 Arduino Uno board
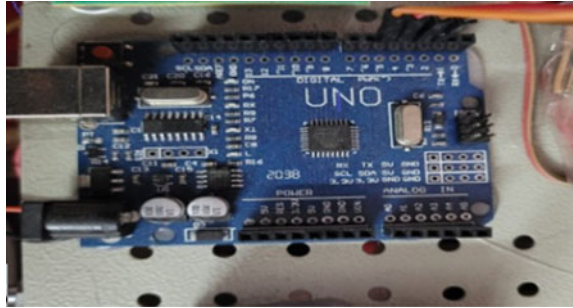
I/O pins (Digital I/O pins(14), analog pins(6)). It is used to control the operation of various machines and devices according to the program given to the ROM of the microcontroller.

3.3 Power Supply

The circuit requires 5 V DC which is obtained using IC 7805 a positive 5 V regulated IC. Power from IC mains is step down to 12 V DC using a suitable regulator circuit, this is given as the input to the regulator IC 7805 whose output is used as supply to the various blocks in the circuit.

3.4 Switches and DC Motor

Figure 3a depicts the picture of limit switch. This switch operates depending upon the motion of machine part or the presence of an object. Figure 3b shows the picture of DC motors. Two dc motors are incorporated in the proposed project for movement of a vehicle in the direction of the signal received.
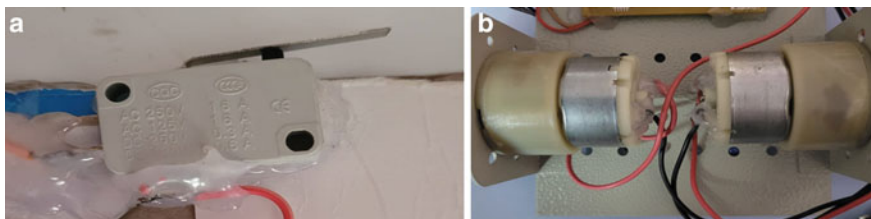
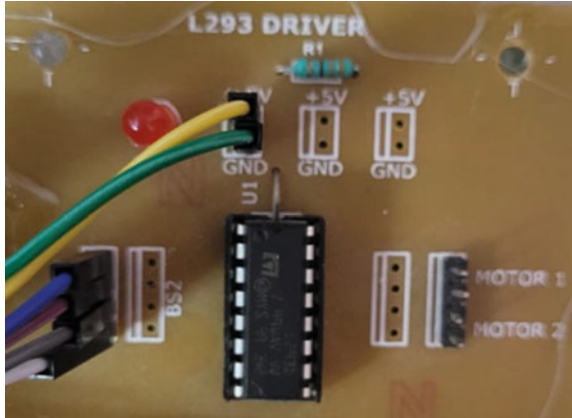**Fig. 3** **a** Switch **b** DC motors

Fig. 4 RF module
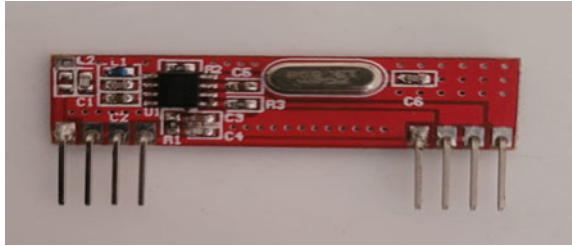
Fig. 5 L293 motor driver IC

3.5　RF Module

Figure 4 shows the picture of RF Module. This module is needed to transmit/receive RF signals between devices wirelessly.

3.6　Motor Driver IC

Figure 5 shows the picture of L293D motor driver IC. It operates on the principle of H-bridge allowing bi directional movement of motor. This driver IC controls two 5 V DC Motors at a time.

4　Results and Discussion

Neck movement controlled vehicle driving system has a transmitter and receiver circuit to assist geriatric as shown in Fig. 6. The system consists of MCU and RF module. The components are classified into two types, one is hardware and the

second one is software components. The neckband has four end stop switches fitted to the RF transmitter, powered by a 9 V battery. RF receiver accepts the commands generated by the neckband. The decoded RF data is given to the microcontroller. This microcontroller drives the DC motor through the H-bridge. Depending on neck movement the vehicle direction is made to move toward Left/Right/Back. The change in direction of the head and the signal is given to the microcontroller. Depending on the direction of the acceleration, the microcontroller controls the vehicle direction like left, right, and back.

The various stages of operation are shown in Figs. 7, 8, 9, 10, 11 and 12. The operation is presented with neck gesture and the controling the moment of vehicle direction is depicted in the following steps.

(i) **Start Button**: This is to initiate the vehicle to receive the command for its operation. It is a switch on the vehicle that makes power supply ON for circuit operation (Fig. 7). It is now ready to receive commands from the neck band to perform four different operations as discussed in following sections.

(ii) **Left Movement**: In the neckband left switch is present. When we slide our neck towards the left side the vehicle will move in the left side direction (Fig. 9).

(iii) **Right Movement**: In the neckband right switch is present. When we slide our neck towards the right side the vehicle will move in the right side direction (Fig. 8).

(iv) **Forward Movement**: In the neckband forward switch is present. When we bend our neck towards down then the vehicle will move in the forward direction (Fig. 10).
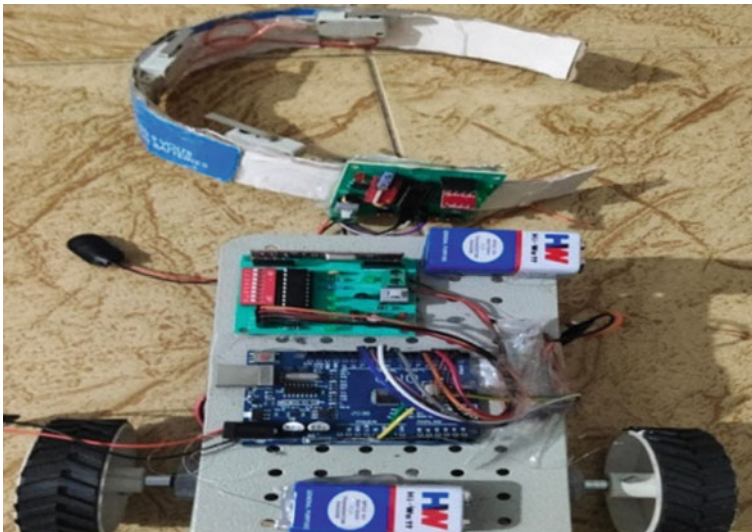

Fig. 6 Prototype of a proposed project

Fig. 7 Neck start movement

Fig. 8 Neck right movement

(v) **Backward Movement**: In the neckband backward switch is present. When we slide our neck towards upside then the vehicle will move in the backward direction (Fig. 11).

(vi) **Stop Button**: This is to bring the vehicle to Halt. It is a switch on the vehicle that turns power supply OFF and stops receiving any commands (Fig. 12).

Fig. 9 Neck left movement

Fig. 10 Neck forward movement

Fig. 11 Neck backward movement

Fig. 12 Neck stop movement

5 Conclusions

The developed module is a prototype to test the functionality of a neck gesture controlled vehicle movement and direction control system. The band on the neck with embedded switches function as a transmitter to control the direction of movement while the Robot module with suitable sensors functions as a receiver, a vehicle that is controlled by the transmitter band on the neck. The vehicle is activated by an ON/OFF switch that makes the vehicle to receive commands for gesture related. Implementing this module in vehicles will add feather to automobile industry in building an Indian low cost and effective vehicle as compared to a similar imported vehicle.

References

1. Megalingam RK, Prakhya SM, Nair RN, Mohan M (2011) Unconventional indoor navigation: gesture-based wheelchair control. In: International conference on indoor positioning and indoor navigation (IPIN), pp 1–4
2. Kanuri S, Janardhana Rao TV, Sridevi CH, Madhan Mohan MS (2012) Voice and gesture-based electric-powered wheelchair using ARM. Int J Res Comput Commun Technol—IJRCCT 1(6):2375–2380
3. Wang D, Yu H (2017) Development of the control system of voice-operated wheelchair with multi-posture characteristics. In: 2nd Asia-Pacific conference on intelligent robot systems (ACIRS), pp 151–155
4. Sundara Siva Kumar V, Ramesh G, Nagesh P (2015) MEMS-based hand gesture wheel chair movement control for disable persons. Int J Curr Eng Technol 5(3):1774–1776
5. Srishti PJ, Shalu SS (2015) Design and development of smart wheelchair using voice recognition and head gesture control system. Int J Adv Res Electric Electron Instrum Eng 4(5):4790–4798
6. Megalingam RK, Chako C (2016) Gesture controlled wheel chair using IR-LED-TSOP pairs along with collision avoidance. In: International conference on robotics and automation for humanitarian applications (RAHA), pp 1–7
7. Ahmed SM, Yasmeen A, Jagadeesh Babu B (2018) Wheelchair with auto navigation for adults with physio and cognitive impairments. Int J Innovative Technol Exploring Eng (IJITEE) 8(252):2278–307
8. Ahmed SM, Shireen A, Jagadeesh Babu B, Shruthi (2019) Powered wheelchair for mobility with features to address physical strength, cognitive response, and motor action development issues. In: International conference on data science, machine learning and applications (ICDSMLA), vol 601, pp 1110–1117
9. Anusha A, Ahmed SM (2017) Vehicle tracking and monitoring system to enhance the safety and security driving using IoT. In: IEEE conference on recent trends in electrical, electronics and computing technologies (ICRTEECT), pp 49–53
10. Braga RA, Petry M, Reis LP, Moreira AP (2011) Intel wheels: modular development platform for intelligent wheelchairs. J Rehabil Res Dev 48(9):1061–1076
11. Simpson RC (2005) Smart wheelchairs: a literature review. J Rehabil Res Dev 42(4):423–436
12. Simpson RC, Levine SP (2002) Voice control of a powered wheelchair. IEEE Trans Neural Syst Rehabil Eng 10(2)
13. Madarasz RL, Heiny LC, Cromp RF, Mazur NM (1986) The design of an autonomous vehicle for the disabled. IEEE J Robot Autom 2(3):117–126
14. Ding D, Cooper RA (2006) Electric powered wheelchairs. IEEE Control Syst Mag 22–34
15. Bourhis G, Moumen K, Pino P, Rohmer S, Pruski A (1993) Assisted navigation for a powered wheelchair. Systems engineering in the service of humans: proceedings of the IEEE international conference on systems man and cybernetics 3:553–558

Improved Numerical Weather Prediction Using IoT and Machine Learning

Rajeshwarrao Arabelli, Chinthireddy Sindhu, Mandadapu Keerthana, Thumma Venu Madhav, Chaganti Vamshi, and Syed Musthak Ahmed

Abstract Weather forecasting is very important for the primary sector to predict the weather conditions particularly for farming. As climate is continuously changing, the older prediction methods have become less effective. To overcome these problems, an improved and more reliable forecasting prediction method is developed. Logistic Regression, Decision Trees, Random Forests, and Support Vector Machine algorithms are used to perform the present study. Here, rain forecasting devices like DHT11 and BMP180 sensors are incorporated. The ESP8266 module is used to add the sensor data to a Firebase database. A current dataset for training and real-time data for testing is utilized in the system. Data mining techniques were used primarily to clean the dataset. Later, the accuracy and time taken to run all the algorithms are calculated. After analyzing all the algorithms, Random Forest algorithm is selected for predicting rain as it gives improved accuracy and reduced execution time to run the algorithm.

Keywords Libraries · Algorithms · Sensors · Firebase

1 Introduction

The use of present technologies for weather prediction, i.e., to forecast the status of the weather, is very important for prediction in the coming days. Meteorologists collect more data about the present situation of the atmosphere in result to know how and when it will evolve in upcoming days. The main atmospheric prediction parameters are air pressure, temperature, wind speed, wind direction, humidity, and precipitation. The data from the observations is integrated during the data assimilation phase with the most up-to-date forecasting methodology. Observations were made. Various weather forecasting systems, spanning from real time to annual climate forecasts, have been created by a number of academics and organizations. Statistical models make up the majority of weather forecasting models. As technology and

R. Arabelli (✉) · C. Sindhu · M. Keerthana · T. V. Madhav · C. Vamshi · S. M. Ahmed
Department of ECE, SR Engineering College, Telangana, Warangal, India
e-mail: rajeshwarrao432@gmail.com

competitiveness advance, many businesses and organizations are turning to data mining approaches for weather forecasting. Data cleaning techniques include LR, Decision Trees, Random Forest, Naive Bayes, and KNN. Both the statistical and mathematical models are static. Data mining may unearth fascinating and instructive patterns from hidden patterns without any prior knowledge. This depends on the data as well as the technique used.

2 Existing Methods

In paper [1], the author employed data mining techniques to forecast weather using a Decision Tree algorithm. In their work, they used the characteristics of maximum and minimum temperature, humidity, and wind speed and obtained an accuracy of 82.62%. Radhika and Shashi [2] in their work used Jupyter Notebook and Machine Learning Algorithm. The information request was created by dissecting time management data of daily weather at a location in order to predict the high temperature of a particular day at the location based on everyday high weather for a span of prior *m* days. Few of their characteristics they used in their experiment were maximum and minimum temperatures, rainfall, cloud conditions, and wind speed. They obtained an accuracy 80%. In [3], NodeMCU, Jupyter Notebook, DHT11, ThingSpeak, and LDR were employed to carry out real-time climate forecast. In [4], Jupyter Notebook and Machine Learning Algorithm were utilized to conduct the experiment by considering the parameters like wind, temperature, wind speed, and wind direction. They attained accuracy of 52.82% in KNN, 42.43% in Naive Bayes, and 82% in the Decision Tree algorithm. Temperature, wind speed, and wind direction are some of the parameters they used in this experiment [4]. In this paper, they proposed weather prediction using Machine Learning and IOT. Arduino, humidity sensor, Machine Learning, and temperature sensor were employed as tools and technologies. Using Machine Learning, this study presents the way to forecast weather conditions and to predict rainfall. The suggested setup will compare predicted values to real-time data and forecast rainfall using the dataset given to the Machine Learning Algorithm [5]. In this paper, they proposed Machine Learning applied to weather forecasting. Jupyter Notebook is one of the gear and technology used in this project. The scope of this paper became confined to predicting the most and lowest temperatures for the subsequent seven days by using climate records from the preceding days [6]. Proposed IoT primarily which is a base on totally ML methods for weather forecast analysis. ML, IoT, DT, SVM, Time Series, Raspberry Pi, temperature and moisture sensor, rain sensor, pressure sensor, and Wireless Access Adapter are a number of the equipment and strategies utilized on this 4 project [7]. ARIMA, IoT, Machine Learning, Time Series Analysis, and the DHT11 temperature and humidity sensor are among the tools and technologies used in this research. Pressure sensor BMP280. The goal of our research is to use an IoT-based smart system to monitor numerous characteristics of weather and anticipate future values [8]. In this paper, they proposed weather prediction for Indian places with the ML methods. LR and ML are some of the tools

and technology employed. Using data science and ML approaches such as linearReg and logisticReg, a simulated system is constructed in this work for determining many climate conditions over the Indian subcountries. Already present meteorological conditions, like temperature, are utilized to fit algorithm, and upcoming changes in the features will be examined using ML techniques [9]. In this paper, Arduino Uno, rain stage sensor, temperature and humidity sensor, soil moisture sensor, WiFi Module, and LCD display are the gear and technology used on this project. The Weather Monitoring and Reporting System project, 5 that is primarily based totally at the Internet of Things, is used to accumulate real-time climate reports. Temperature, humidity, and moisture will all be monitored [10]. In [11], the mix of Naïve Bayes and Chi-Square algorithm to predict weather condition is presented. In their work, the constant information, i.e., time-series data, is assembled and analyzed and this dataset is used as an interface for Weather Prediction System. The use of data mining techniques in forecasting weather is analyzed in [12, 13]. They proposed a service-oriented architecture to forecast to collect weather information using data mining technique. In [14], the weather parameters are predicted using ARIMA model using R studio where the prediction data is collected every month and every year. In their work, the data is collected from month of June. Only first two successive years give accurate results in prediction of future weather values in their project. A simple model for weather forecasting has been described in [15]. In their model, they made use of simple mathematical equation using Multiple Linear Regression (MLR) equations which can be easily understood by farmers. In their work, the data of a particular station is recorded which is a time-series data. In their work, they made use of weather parameters such as minimum and maximum temperature and relative humidity as prediction parameters.

3   Proposed Method

We have selected the method which is to create a real-time weather forecasting system using a technique. The system employs a temperature and humidity sensor, the DHT11, and a BMP280 pressure sensor. The sensor data is uploaded to a Firebase database using the ESP8266 module. We will employ Logistic Regression, Decision Trees, Random Forests, and Support Vector Machine in this project. We will use an existing dataset for training and real-time data for testing. To begin, we will clean the dataset using data mining techniques. We will calculate the accuracy of all of the specified methods, then train them using current training data, and use them to forecast the weather using real-time testing data.

Figure 1 employs temperature and humidity sensor, DHT11, and BMP280 pressure sensor. From these sensors, we will get temperature, humidity, and pressure values. These values are uploaded to a Firebase database using the ESP8266 module. ESP8266 module acts as intermediate element between sensors and Firebase database. The data is then exported as JSON file. JSON file is then converted to CSV file from Firebase. This data acts as a testing data for ML model to predict rain. SVM,
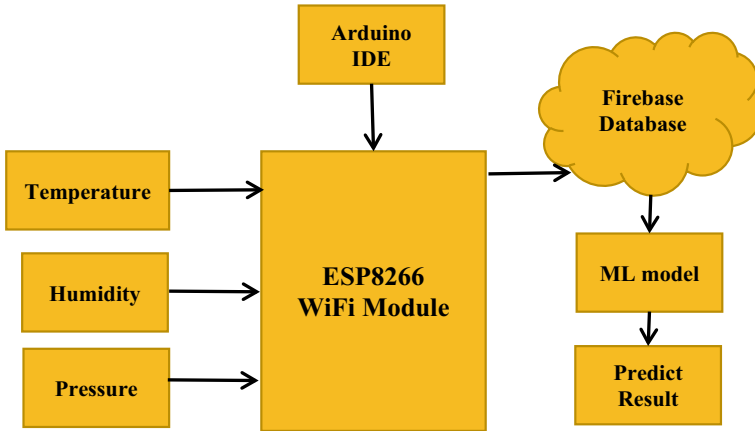
Fig. 1 Block diagram of proposed system

Decision Trees, Logistic Regression, and Random Forests were employed. Existing dataset is taken from Kaggle. Data cleaning, data pre-processing, and feature selection are done on dataset. All specified Machine Learning models were applied on dataset to find accuracy as well as time taken to run that algorithm. Random Forest gives little more accuracy, and it took less time to run. The model was trained using existing training data, and weather rain will fall or not was predicted using testing data.

4 Implementation of Proposed System

The complete execution of the project is explained by a flowchart as shown in Fig. 2.

The hardware requirements are DHT11 sensor, BMP180 sensor, and ESP8266 sensor. Both DHT11 sensor and BMP180 sensor are connected to ESP8266 powered by USB cable. The connectivity of the DHT11 sensor and BMP180 sensor with ESP8266 is given in Table 1.

Google Colab software is used in the present work. The python code is written in Google Colab software. All the algorithms, namely SVM, Decision Trees, Logistic Regression, and Random Forests, were implemented in the software.

5 Methodology

The following steps are followed in developing the project

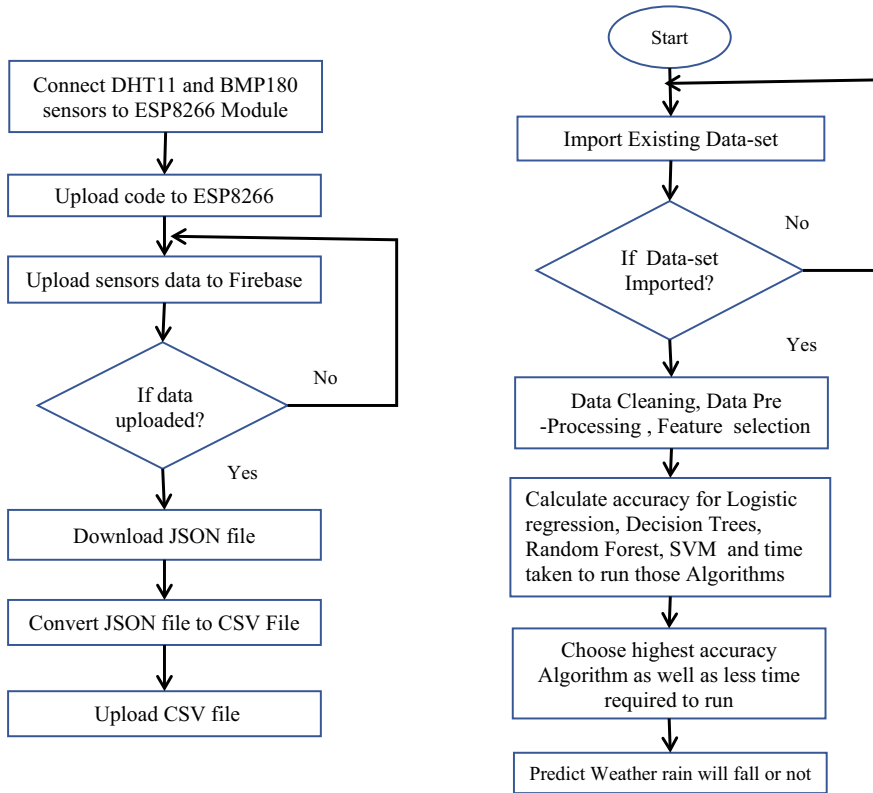1. Performed data collection and pre-processing.

Fig. 2 Flowchart of proposed system

Table 1 Pin configuration of DHT11 and BMP180 with ESP8266

DHT11 pin connections with ESP8266		BMP180 pin connections with ESP8266	
VCC	3V3	GND	GND
Data	D5	VIN	3V3
GND	GND	SDA	D2
		SCL	D1

2. Imported the required libraries such as Numpy, Pandas, sklearn, and scipy.
3. Uploaded the dataset and removed the maximum null values rows and unwanted rows as well as removed outliers.
4. Transformed the categorical columns to numerical columns and pre-processed data.
5. Calculated the accuracy and time taken for each algorithm.
6. Selected Random Forest algorithm which has high accuracy and taken less time to run the algorithm.

7. Calculated real-time testing data of temperature, humidity, and pressure and stored in Firebase database.
8. Predicted weather rain will fall or not using real-time values.

6 Results and Discussion

The training data is collected from Kaggle website which includes weather parameters such as temperature, wind speed, wind direction, humidity, and pressure. Prior to this, data pre-processing and data cleaning are performed. Later, accuracy and execution time for each algorithm are calculated. This is listed in Table 2.

From the above comparison Table 2, it is observed that Random Forest algorithm is better in terms of accuracy and execution time. Figure 3 gives the hardware implementation of proposed system to get real-time data. The DHT11 sensor and BMP180 sensor are connected with ESP8266 pins, and power supply is provided by USB cable to connect it to the system.

Figure 4 gives the Firebase output of the proposed system. The real-time values sensed from DHT11 and BMP180 pressure sensor are stored in Firebase database using ESP8266 WiFi module. The values stored in the Firebase database are exported from database as JSON file. This is shown in Fig. 5.

Table 2 Accuracy and time taken by different algorithms

Algorithm	Accuracy (%)	Time taken (s)
Logistic regression	83.9	0.26
Decision trees	76	0.26
Random forest	84.2	2.7
Support vector Machine	83.6	96.6
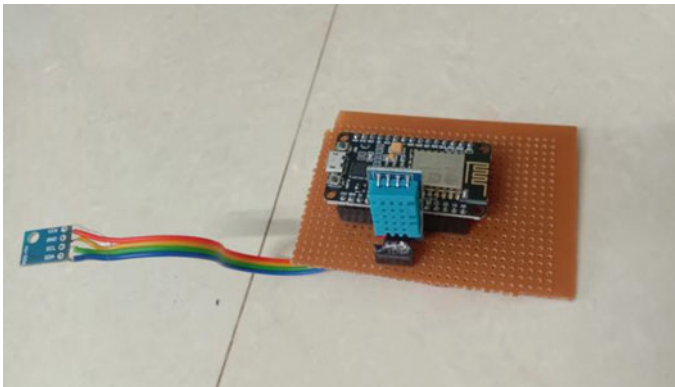
Fig. 3 Hardware of the proposed system
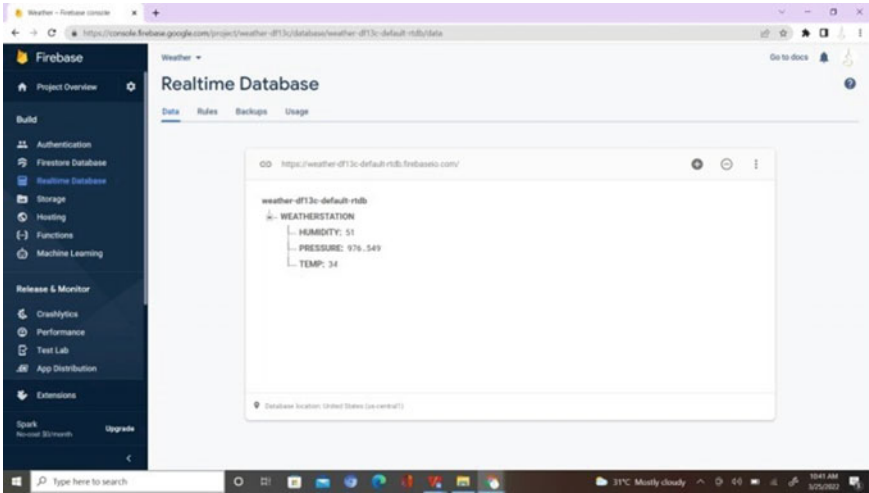
Fig. 4 Firebase database output

Fig. 5 JSON file output

Next, JSON file is converted into CSV file and then uploaded in Google Colab to predict the rain as shown in Figs. 6 and 7.

The sensor data is sent to Firebase database. This data is converted to CSV file. This data acts as testing data and is applied with ML model to find weather rain will fall or not on a particular day. The existing dataset is taken from Kaggle website. In order to predict the rain, we need to import certain libraries, namely Numpy to

Fig. 6 CSV file uploaded in Google Colab

```
if round(qw[0])==0:
  print("Rain will not fall")
else:
  print("Rain will fall today")


Size of weather data frame is : (1, 3)
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:
  del sys.path[0]
[0.0]
Rain will not fall
```

Fig.7 Output of random forest algorithm

perform mathematical operations, Pandas to load dataset to program, scipy to remove outliers, sklearn to perform data pre-processing, and feature selection to import ML algorithms and to find accuracy of each model. Z-score was performed to detect and remove the outliers in dataset using scipy library. The categorical columns were transformed to numerical columns, and the data was pre-processed. In data pre-processing, the entire values in the respective column are transformed in between the range of 0 and 1 so that operations are performed easily and we get accurate output. The top independent columns were selected using SelectKBest method in sklearn library. After performing all this data is updated. The updated data is fed to specified ML algorithms. We have taken 4 algorithms, namely SVM, Decision Trees, Logistic Regression, and Random Forests. We calculated the accuracy and time taken to run the algorithm. The below table shows the accuracy and time taken for each algorithm.

7 Conclusion

In the present work, the real-time weather forecasting parameters are built using low-cost sensors. The three most critical characteristics that are monitored and posted on Firebase cloud are temperature, pressure, and humidity. The Google Colab environment is incorporated to predict the rain by using Decision Trees, Logistic Regression, Random Forests, and SVM. By incorporating this algorithm, accuracy and time taken to run the code are calculated. After analyzing the various models, it is found that Random Forest algorithm has proven good results with an accuracy of 84% with a duration of 2.4 s. Results have been obtained by incorporating all the four algorithms in the present work as compared to the previous works where each algorithm is implemented independently. The system may also be adapted for commercial use, and it has a variety of applications in smart homes, buildings, sports, and health care, to name a few.

References

1. Bhatkande1 SS, Roopa G, Hubballi (2016) Weather prediction based on decision tree algorithm using data mining techniques. Int J Adv Res Comput Commun. Eng. 5(5):484–487
2. Radhika Y, Shashi M (2009) Atmospheric temperature prediction using support vector machines. Int J Comput Theory Eng 1(1):55–58
3. Verma G, Mittal P, Farheen S (2020) Real time weather prediction system using IOT and machine learning. Int Conf Signal Process Commun 322–324
4. Khan ZU, Hayat M (2014) Hourly based climate prediction using data mining techniques. Middle-East J Sci Res 21(8):1295–1300
5. Gopinath N, Vinodh S, Prashanth P, Jayasuriya A, Deasione S (2020) Weather prediction using machine learning and IOT. Int J Eng Adv Technol 9(4):2249–8958
6. Holmstrom M, Liu D, Vo C (2016) Machine learning applied to weather forecasting stanford University 15 Jan 2016
7. Nallakaruppan MK, Senthil Kumaran U (2019) IoT based machine learning techniques for climate predictive analysis. Int J Recent Technol Eng 7(5S2):2277–3878
8. Raval MP, Bharmal SR, Hitawla FAA, Gupta P (2020) Machine learning for weather prediction and forecasting for local weather station using IoT. Int Res J Eng Technol 07(02):419–423
9. Shivang J, Sridhar SS (2018) Weather prediction for Indian location using machine learning. Int J Pure Appl Math 118(22):1314–3395
10. Bhagat AM, Thakare AG, Molke KA, Muneshwar NS, Choudhary V (2019) IOT based weather monitoring and reporting system project. Int J Trend Sci. Res. Develop. 3(3):2456–6470
11. Biswas M, Dhoom T, Barua S (2018) Weather forecast prediction: an integrated approach for analyzing and measuring weather data. Int J Comput Appl 182(34):20–24
12. Wang ZJ, Mazharul Mujib ABM (2017) The weather forecast using data mining research based on cloud computing. J Phys Conf Ser 91(1):1742–6596
13. Chen N, Qian Z, Nabney IT, Meng X (2014) Wind power forecasts using gaussian processes and numerical weather prediction. IEEE Trans Power Syst 29(2)
14. Kothapalli S, Totad S (2017) A real-time weather forecasting and analysis. IEEE Int Conf Power Control Signals Instrum Eng 1567–1570
15. Paras SM (2016) A simple weather forecasting model using mathematical regression. Environ Sci Indian Res J Extension Educ 16 Jan 2016

Inductive Coupling-Based Wireless Power Transmission System in Near Field to Control Low-Power Home Appliances

Srinivas Samala, M. Srinayani, M. Rishika, T. Preethika, K. Navaneeth, G. Nandini, and Syed Musthak Ahmed

Abstract An intimidating and challenging topic is the wireless delivery of electrical power to a load. Wirelines are currently being utilized to transfer power, but there are losses associated with this method. The invention of wireless power transmission (WPT) has opened up a new world to us. A magnetic field is used to transmit power from a transmitter coil to a receiver coil in WPT utilizing inductive coupling, which is part of NFWPT. Home appliances and the healthcare industry are embracing inductive coupling as an efficient method of transferring electricity over short distances. Wireless power transfer is also boosting the popularity of electric vehicle charging stations, biomedical implants, consumer electronics, industrial applications, etc. This paper discusses low-cost prototype of WPT based on mutual coupling between two coils. Here in the proposed system, two copper windings of fixed turns and SWG are placed face two face on the same axis. Due to mutual coupling, the electrical energy is transform from the primary to secondary coils. Experimentation is carried out by varying the gap between the coils, and the coupled power at the receiving end is utilized to run a low-power house appliance.

Keywords Wireless transmission · Inductive coupling · Near-field communication

1 Introduction

There is a growing industry in which electrical power can be supplied without the requirement of interconnecting cables by transferring energy wirelessly. Wireless power transfer (WPT) using inductive coupling sends magnetic field power to a receiver coil. Wireless low-power transfer (WLPT) is a need for charging smart phones, electric vehicles, and other electrical gadgets. Medical implanted devices, in particular, have benefited greatly from the use of WPT [1]. The wireless power transmission is possible in two ways; they are far field wireless power transmission (FFWPT) and near field wireless power transmission (NFWPT) [2]. Magnetic

S. Samala (✉) · M. Srinayani · M. Rishika · T. Preethika · K. Navaneeth · G. Nandini · S. M. Ahmed
S R Engineering College, Warangal, India
e-mail: srinu486@gmail.com

coupling between two coils restricts the transmission range of the NFWPT. However, FFWPT transmission supports long-range transmission, but it is less efficient than near field transmissions.

Despite the low mutual inductance between coils, increasing power transfer while maintaining system efficiency was the major objective of WPT research. Transmission of electricity from the transmitter to the receiver is accomplished by transferring magnetic fields between two resonant circuits, known as a resonant inductive coupling [3, 4]. It is possible to transmit more power over longer distances by making use of the weaker magnetic fields in outer areas ("tails") of near fields by means of resonance. High efficiency can be achieved at a range of 4–10 times the coil diameter using resonant inductive coupling [5]. One of the factors affecting the efficacy of inductive coupling is the decrease in dipole fields between the transmitter and receivers; this is related to the cube of the transmitter–receiver distance. Hence, shorter distances achieve higher efficiency [6].

Wireless power system can be achieved by inductive or magnetic coupling between two coils. The primary winding of the WPT is used to link the power supply, and the secondary winding is used to connect the load. In [7], their work designed to achieve extremely efficient transmission, and they merged an AC/DC converter together with an antenna system with ideal coupling coefficients into a wireless power transmission system employing magnetic coupling. In [8], the authors discussed wireless power transmission system with high-frequency resonant inductive coupling. The HFWPT system is intended to operate at a fifty kilohertz resonance frequency to provide better efficiency. Several device networks and cellular networks use identical wireless transmission technique to improve their transmission efficiency. In [9], the authors are addressed issues related to high-resonance system. The frequency variation of resonance is restricted, and the effectiveness of the energy exchange varies supported the coupling magnitude in this work. In [10], the authors examined the viability of using non-radioactive fields at the ends of resonant objects for intermediate energy transfer. Intuitively, 2 resonant things with an equivalent resonant frequency tend to interchange energy effectively, whereas superfluous off resonant objects dissipate relatively very little energy. In [11], the authors designed a basic wireless power transfer system for smart home applications. An electrodynamic induction approach was used to create this apparatus to show wireless power transfer. The stability analysis of an ICPT system coupled to an electric bicycle was addressed in [12]. Magnetic resonance coupling to create a wireless power transfer system that is both efficient and powerful is presented in [13]. An antenna system's equivalent circuit and simulation were used to derive the theoretical method to optimize antenna performance [14].

2 Existing Systems

Several works are reported in the design and development of wireless power transmission systems. The basic principle behind the WPT is inductive or magnetic coupling between two coils. A block diagram of this type of system is depicted in Fig. 1.

Power transmission system based on electromagnetic coupling circuit is illustrated in Fig. 1. The coupling model was used to investigate ideal coupling coefficients in terms of mathematical expressions. Resonance frequency is critical to system design because magnetic resonance coupling entails producing a resonance and transmitting power without the emission of electromagnetic waves. The various design parameters are

$$\omega_0 = \frac{1}{\sqrt{LC}} \tag{1}$$

$$Q = \sqrt{\frac{L}{C}}\frac{1}{R} = \frac{\omega_0 L}{R} \tag{2}$$

Coil, capacitor, and resistor are all components of an electromagnetic resonance circuit depicted in Fig. 2. Energy in this circuit oscillates at a specific resonance frequency between the coil and the capacitors, which store the energy in both a magnetic and an electric field. Circuit losses are reduced as a result of an increase in the system's quality factor (the reduction of R) [15, 16].
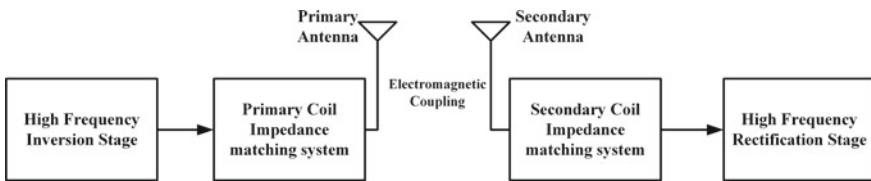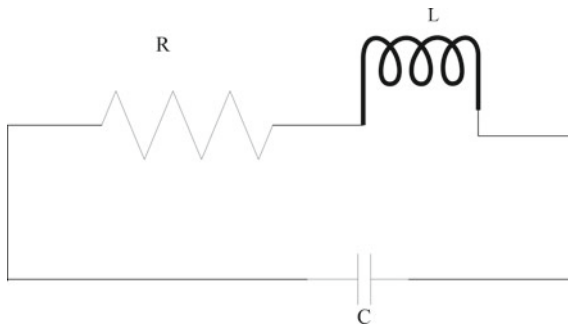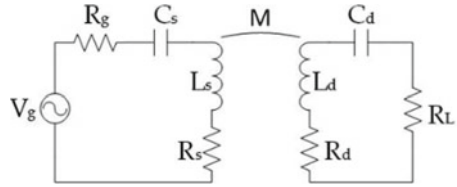

Fig. 1 WPT system's diagram

Fig. 2 Resonator circuit

Fig. 3 Coupled resonator

It is essential that a high-resonance WPT system has an efficient mechanism to transfer energy. A high-quality factor electromagnetic resonator is often manufactured using conductive components with limited resonant frequency widths. Depending on the coupling ratio k and the particular resonator, the energy exchange will be adversely affected [17, 18]. A two-resonator system's dynamics can be identified using coupling modal theory or comparable circuit analysis. Coupled resonators circuit is depicted in Fig. 3.

R_s is the source internal resistance, and the voltage source's amplitude is referred to as V_g. Coils L_m and L_d are the source and the device resonators. The mutual inductance is represented by $M = k\sqrt{L_m L_d} * E$. A resonator is created by connecting a capacitor in series with each coil. Resistors R_s and R_d denote the ohmic and radioactive losses in each resonator's coil; R_L represents the AC load resistance. It is possible to calculate the yield of this circuit by comparing to the maximum power available from the source, the amount of energy delivered for resistance to loads [19, 20]. The expression for ratio is given in Eq. 4.

$$\frac{P_L}{P_{g,\max}} = \frac{4U^2 \frac{R_g}{R_s} \frac{R_L}{R_d}}{\left[\left(1 + \frac{R_g}{R_s}\right)\left(1 + \frac{R_L}{R_d}\right) + U^2\right]^2} \tag{3}$$

$$U = \frac{\omega M}{\sqrt{R_s R_d}} = k\sqrt{Q_s Q_d} \tag{4}$$

The optimal system performance can be achieved by selecting the right load and source resistances or by capturing alternative resistance values via an impedance matching connection. Equation 5 represents the efficiency of power transfer.

$$\eta_{\text{opt}} = \frac{U^2}{\left(1 + \sqrt{1 + U^2}\right)^2} \tag{5}$$

Systems with large U values are capable of transferring energy with high efficacy. Finding the system's performance factors, such as U, Q_s, and Q_d can help to maximize system's efficiency. An electric toothbrush is an example of an induction-based wireless power transmission system that uses high coupling values and a short range. Traditional induction systems are less efficient than high-quality resonators. In addition, low coupling values can now be achieved more effectively. It is not necessary to precisely place the source and device because of this. High value of Q maximizes

the peak voltage of capacitor. Equation 6 describes the connection between capacitor voltage peak and quality.

$$V_{C_{\text{peak}}} = Q\frac{2V_s}{\pi} \tag{6}$$

3 Proposed System

Figure 4 depicts the proposed system's block diagram. It consists of the primary coil associated circuit and the secondary coil associated circuit, medium between the two coils is air. Mutual coupling of two coils determines the output voltage to drive the low-power appliance. The mutual coupling L_m is proportional to the product of the primary and the secondary coil inductance (L_p, L_s).

Here, power input is given to the primary coil via heat sink, the secondary coil is connected to the low-power home appliance via a storage device, and the voltage to the device is stabilized through a Zener diode across the capacitor. Figure 5 depicts the system that was built.
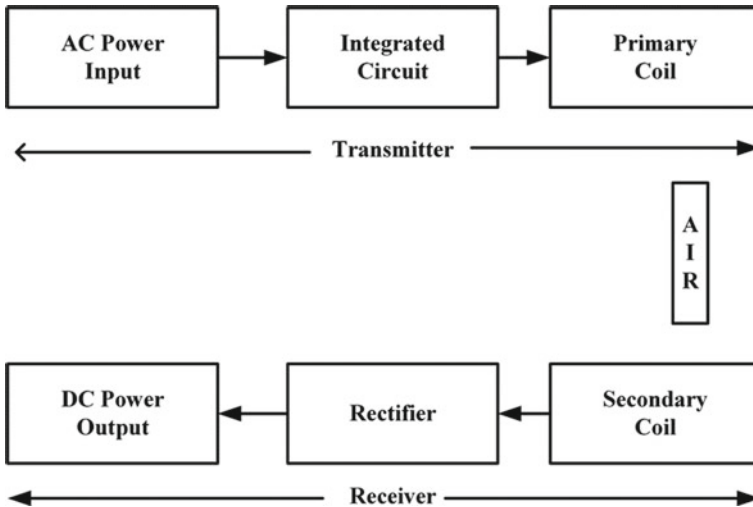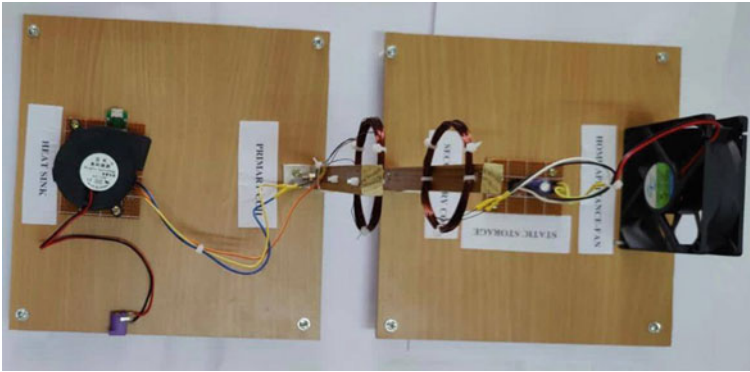

Fig. 4 Proposed work model

Fig. 5 Prototype of wireless power transmission system

4 Results and Discussion

The setup of wireless power transmission system is shown in Fig. 5. The primary and secondary setups for experimentation are shown in Fig. 6. Figure 6.a shows the primary side of the setup which contains power supply mains, a heat sink, and coil, while Fig. 6.b shows the secondary side of the setup containing a storage device connected to an appliance.

The coupling between the coils by variation of gap between the coils is understood in Fig. 7.

As the gap between the primary and secondary coils was varied, there was a drop in the voltage and current due to the change in the mutual coupling between the coils. This is being shown in Table 1.

The input and output power obtained are tabulated and simulated using MATLAB to obtained the efficiency of the system. This is shown in Fig. 8. The efficiency of the system reduces as the coupling distance increases. By fixing an appropriate distance between the coils, the required power to run an appliance can be achieved. With
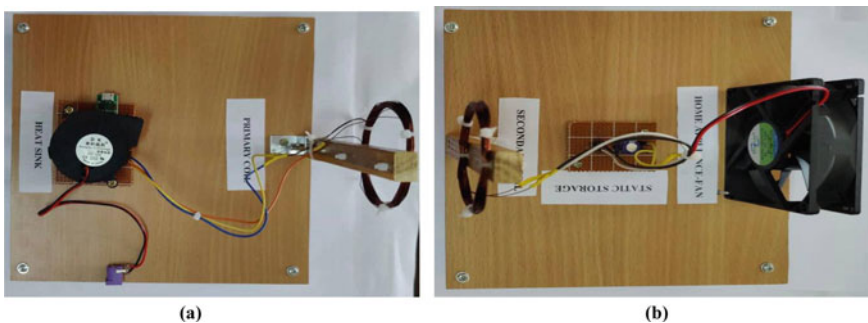

Fig. 6 a Primary coil setup **b** secondary coil setup
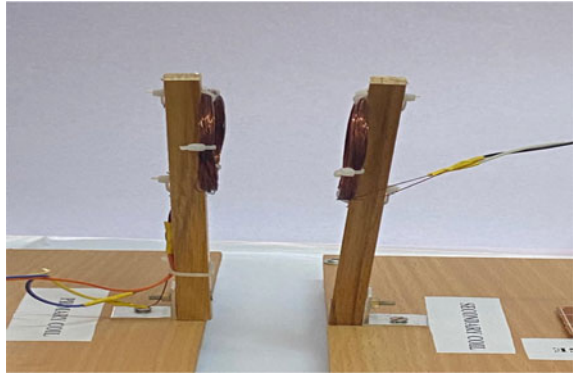
Fig. 7 Primary and secondary coupling

Table 1 Practical values of voltage and current recorded while the coils gap is varied

Distance (in cm)	V_{in} (v)	I_{in} (A)	P_{in} (W)	V_{out} (v)	I_{out} (A)	P_{out} (W)	% η
0	212	0.15	31.8	23.9	0.85	20.31	63.86
0.5	212	0.15	31.8	20.1	0.81	16.28	51.19
1	212	0.15	31.8	14.8	0.72	10.65	33.49
1.5	212	0.15	31.8	10.3	0.59	6.077	19.11
2	212	0.15	31.8	5.2	0.32	1.66	5.22
2.5	212	0.15	31.8	3.1	0.2	0.62	1.94
3	212	0.15	31.8	1.2	0.06	0.072	0.22

the present product developed, a 12 v mini fan is successfully run to check the functionality.

5 Conclusion

In the present work, a wireless power transmission system is developed that can run an home appliance. The principle of inductive coupling is to generate power wirelessly which is demonstrated, and it is further investigated the transfer of electricity wirelessly over a distance of 3 cm between the primary and secondary coils. Heat sink is used in the circuit to protect IC from power dissipation; from the simulation results, it is observed that as distance between the coils increases, it effect the efficiency of the system. Since the results of this study are sufficient to power the home appliance, so wireless power transfer can be regarded a viable option in a wide variety of other applications.
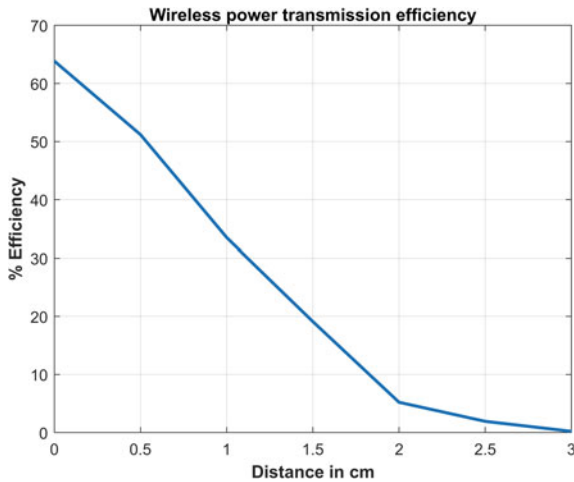
Fig. 8 Simulation results showing efficiency of WPT versus coil gap/power variations

References

1. Kesler MP (2013) Highly resonant wireless power transfer: safe, efficient, and over distance. http://large.stanford.edu/courses/2016/ph240/surakitbovorn1/docs/kesler.pdf. Accessed 14 Nov 2020
2. Hassan MA, Elzawawi A (2015) Wireless power transfer through inductive coupling. Recent Adv Circ 115–118
3. Van Schuylenbergh K, Puers R (2009) Inductive powering: basic theory and application to biomedical systems. Springer, Netherlands
4. Kuipers J et al. (2012) Near field resonant inductive coupling to power electronic devices dispersed in water. Sens Actuators A: Phys 178(2012):217–222
5. Ahmed M, Ansari MD, Singh N, Gunjan VK, BV SK, Khan M (2022) Rating-based recommender system based on textual reviews using IoT smart devices. Mob Inf Syst
6. Bou Balust E, Alarcón Cot EJ, Gutiérrez Cabello J (2012) A comparison of analytical models for resonant inductive coupling wireless power transfer. In: PIERS 2012: Progress in electromagnetics research symposium: proceedings, Moscow, Russia, The Electromagnetics Academy, pp 19–23
7. Shidujaman M, Samani H, Arif M (2014, May) Wireless power transmission trends. In: 2014 International conference on informatics, electronics & vision (ICIEV), IEEE, pp 1–6
8. Lakshmanna K, Shaik F, Gunjan VK, Singh N, Kumar G, Shafi RM (2022) Perimeter degree technique for the reduction of routing congestion during placement in physical design of VLSI circuits. Complexity 2022, 1–11.
9. Mohammed SS, Ramasamy K, Shanmuganantham T (2010) Wireless power transmission–a next generation power transmission system. Int J Comput Appl 1(13):100–103
10. Agcal A, Ozcira S, Bekiroglu N (2016) Wireless power transfer by using magnetically coupled resonators. J Wirel Power Transf FundamTechnol 49–66
11. Jameela T, Athotha K, Singh N, Gunjan VK, Kahali S (2022) Deep learning and transfer learning for malaria detection. Comput Intell Neurosci, 2022
12. Kim YH, Kang SY, Cheon S, Lee ML, Zyung T (2010, June) Optimization of wireless power transmission through resonant coupling. In: SPEEDAM 2010, IEEE, pp 1069–1073
13. Kumar V, Kumar MR, Shribala N, Singh N, Gunjan VK, Arif M (2022). Dynamic Wavelength Scheduling by Multiobjectives in OBS Networks. J Math, 2022

14. Kurs A, Karalis A, Moffatt R, Joannopoulos JD, Fisher P, Soljacic M (2007) Wireless power transfer via strongly coupled magnetic resonances. Science 317(5834):83–86
15. Senthil Kumar G, Ramana K, Madana Mohana R, Aluvalu R, Gunjan VK, Singh N (2022) An effective bootstrapping framework for web services discovery using trigram approach. Mob Info Syst, 2022, 1–12
16. Wang J, Li J, Ho SL, Fu WN, Li Y, Yu H, Sun M (2012) Lateral and angular misalignments analysis of a new PCB circular spiral resonant wireless charger. IEEE Trans Magn 48(11):4522–4525
17. Cannon BL, Hoburg JF, Stancil DD, Goldstein SC (2009) Magnetic resonant coupling as a potential means for wireless power transfer to multiple small receivers. IEEE Trans Power Electron 24(7):1819–1825
18. Fareq M, Fitra M, Irwanto M, Hasan S, Arinal M (2014, April) Low wireless power transfer using inductive coupling for mobile phone charger. In: Journal of physics: conference series, vol 495(no 1) IOP Publishing, p 012019
19. Imura T, Hori Y (2011) Maximizing air gap and efficiency of magnetic resonant coupling for wireless power transfer using equivalent circuit and Neumann formula. IEEE Trans Industr Electron 58(10):4746–4752
20. Buja G, Bertoluzzo M, Mude KN (2015) Design and experimentation of WPT charger for electric city car. IEEE Trans Industr Electron 62(12):7436–7447

IoT-Based Safety and Security System for House Boats

Rajeshawarrao Arabelli, Nikitha Adepu, Bheemreddy Varshitha, Lethakula Abhinaya, Vasanth, and Syed Musthak Ahmed

Abstract Waterways are one of the means of transport for the people staying in islands and coastal areas. Safety is one of the major factors that have to be considered in waterways. Present work explains the design and creation of an embedded system to detect the obstacle and identify a safety gateway. Here the boat is fixed with an obstacle detection module (ODM) that tracks the presence of obstacle in three directions of the boat moment, viz. front, left, and right. Once the obstacle is detected, the Notification Module (NM) responds by giving a message indicating to halt the boat or to take a deviation away from the obstacle. In the developed system, an RF transmitter and receiver are incorporated. The RF transmitter transmits the signal to the RF receiver during the instance of collision and/or emergency to the control room for necessary action. The developed prototype can be deployed in various types of boats which include passenger vessel, cargo vessel, fish boats, etc.

Keywords Waterways · Ultrasonic sensor · RF module · Arduino UNO

1 Introduction

While travelling in boats, there could be many obstacles during the voyage such as ice burgs, rocks, large fish, etc. The obstacle is detected with Sound Navigation and Ranging (SONAR). Active and passive sonar are the two types of sonar available. Passive sonar takes sound waves from another source and converts them to electrical signals, whereas active sonar puts out sound waves in pulses. Sonar is utilized in a variety of industries, including oil and gas, fishing, mining, and seafloor mapping. In general, accidents are common either on roadways or waterways but the reason behind the accident will be different. The reasons for road accidents are over speed, violating traffic signals, and poor driving skills. Similarly, the reasons behind the accidents in waterways are collision, sudden weather changes, navigation rules violations, machinery failure, and operators' inexperience.

R. Arabelli (✉) · N. Adepu · B. Varshitha · L. Abhinaya · Vasanth · S. M. Ahmed
SR Engineering College, Warangal, Telangana, India
e-mail: rajeshwarrao432@gmail.com

2 Literature Review

This section explains the different existing methods suggested by different authors and they are:

Muhammad Baballe Ahmed and Mehmet Cavas [1] in their work created a reliable cost-effective home security system. They developed a secure system where in any one crossing the wall will be detected by suitable devices incorporated in the system. They also installed a CCTV to capture the real-time video. In [2–4], the authors developed a security detection system to secure any of the walls that were crossed. The author used LDR and LED. LCD is also employed to display the status of the walls and security detection system that will monitor, guide, and protect the environments that needs to be secure against burglars and abductors. Similar works on security system on four wall crossing detection is carried out. In [5], the authors developed a security alarm detection system using Arduino microcontroller to read the dis carried out data once the PIR sensor detects an intruder. The GSM Module is used to receive the instructions from the microcontroller and sends a message to the designated mobile number which is a system that protects someone or something by employing the Internet of Things when the PIR sensor detects an inducer [6, 7].

The authors of [8] presented radar based on an Arduino-compatible ultrasonic sensor. The radar can detect things at a distance of 2 cm to 4 m and display the distance to the object as a value. They've also included a motion sensor that can detect movement from any angle. Lakshnanna et al. [9] presented a system for ensuring the safety and security of people and machines in hazardous environments. They used an ultrasonic sensor and Zigbee to identify objects and deliver emergency notifications. It can monitor remote locations and relay signals wirelessly in this project. In [10], the authors developed a distance measurement technique that is both effective and trustworthy. In order to improve the resolution of time of flight (TOF) measurements, the author also implemented appropriate signal processing techniques. In [11], the authors developed a simple climate and alert system with minimum effort which allows the family members to follow up about issues like unexpected climate changes or means of crisis. In the event, if the anglers cross the boundary, the buzzer begins to ring and the vessel comes to halt. In [12], Sohanur Rahman implemented a system which provides reliable information to make better decisions in order to avoid collisions, fire, overloading, etc., in order to ensure greater safety for passengers [13, 14].

In [15], the authors developed a conceptual model for an overloaded sensor. They used cutting-edge technology that combines elevator concepts with load sensors (HCC-High-Capacity Compression) and batching controllers. In [16], Fei Cheng and Spyros Hirdaris, they have improved the ship safety through many research and innovations. They have classified many rules like strength, propelling machinery, electric system, control system, and these are interlinked with IMO conventions. Based on this, they have provided many safety measures of ships. In [17], the authors have improved the ship safety through many research and innovations. They classified many rules like strength, propelling machinery, electric system, and control system,

and these are interlinked with IMO conventions. Based on this, they have provided many safety measures of ships. They developed a security system in [18] ASA EK to give measures of maritime safety regulatory framework established by the ISM Code, as well as safety culture and safety management in the Swedish national and international shipping business, which includes passenger transport. They also introduced numerous passenger safety measures [19].

3 Proposed System

The power supply is connected to the Arduino and DC motor. Three ultrasonic sensors are connected to the Arduino in three directions, i.e. front, left, and right. The buzzer and push-button are also connected to Arduino for emergency cases. Arduino is connected to an RF wireless transmitter to pass RF signals to the receiver section. We include an emergency help box with an emergency button in this security system. The emergency button can be pressed if the passengers on the boat are in danger. The RF wireless transmitter provides a signal to the RF wireless receiver after pushing an emergency button, indicating the alarm to the boat management team for the rescue process.

The HT 12E encoder is used in the RF transmitter, while the HT 12D decoder is used in the receiver. The block diagram of transmitter and receiver is shown in Fig. 1 while Fig. 2 illustrates the prototype with sensors and indicating devices.
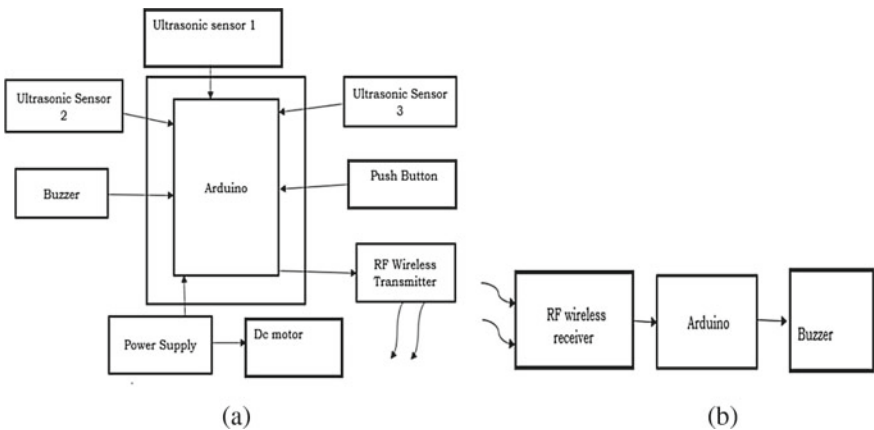

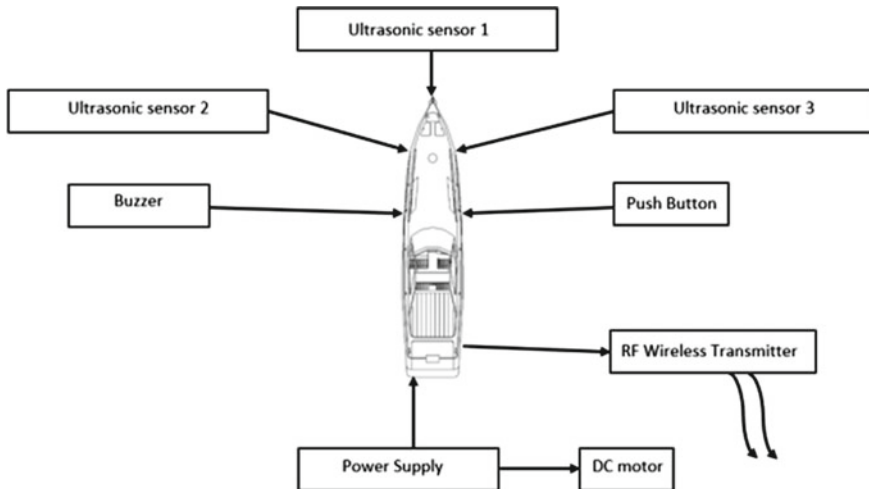**Fig. 1** Block Diagram of **a** transmitter module and **b** receiver module

Fig. 2 Boat with sensors and indicating devices in present implementation

4 Hardware Requirements in the Development of the Product

The following are the major components that are used in developing the proposed model (Fig. 2).

Ultrasonic Sensor

The ultrasonic sensor is a device that uses ultrasonic waves emitted by the sensor to calculate the distance between two objects. The ultrasonic sensor gives information about an object's vicinity by transmitting and receiving ultrasonic pulses through a transducer. The borders are defined by unique echo patterns produced by high-frequency sound waves. Ultrasonic sensors function by transmitting sound waves at a high frequency in comparison with human hearing. To transmit or receive ultrasonic sound waves, the sensor transducer works as a microphone. Proximity sensors are a term used to describe non-contact sensors. Ultrasonics is self-contained in terms of light, smoke, dust, and colour. Target detection at a long distance is done using a variety of surface properties (Fig. 3).

RF Transmitter and Receiver

The designer of wireless systems confronts two significant challenges: it should operate over a specified range of distance and must convey a specific amount of data at a specific rate. The RF modules are small in size and operate at a wide variety of voltages, from 3 to 12 V. 433 MHz is used by the RF transmitter and receiver modules (Fig. 4).
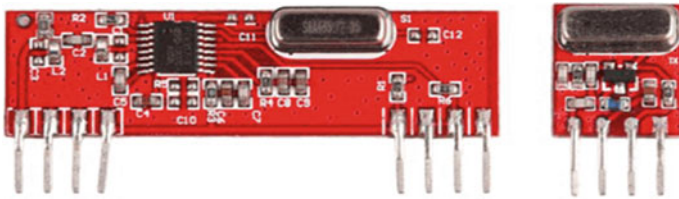
Fig. 3 Real-time ultrasonic sensor

Fig. 4 Real-time RF modulator

Features of RF module contains reception frequency of the RF Module is 433 MHz. The RF module receiver's typical frequency is 105 Dbm. The receiver has a 3.5 mA power supply. The RF module uses very little electricity. The voltage at which the receiver operates is 5 V. The frequency range of an RF module transmitter is 433.92 MHz. An RF module transmitter's supply voltage is 3v ~ 6v. The output power of an RF module transmitter is 4v ~ 12v.

While transmitting good judgement zero and entirely suppressing the provider frequency, the transmitter pulls no energy, costing significantly less electricity in battery operation. With a 3 V energy supply, when common sense one is sent, service is entirely direct to roughly four 5 ma. The data from the transmitter is delivered in serial format to the tuned receiver (Fig. 5).
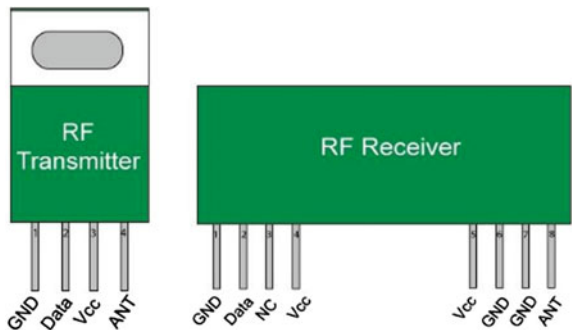

Fig. 5 Basic RF module structure

Fig. 6 Arduino Uno

Arduino UNO

Arduino Uno is a basically 8 bit microcontroller board. It supports 6 analogue pins, 14 digital pins, and various protocols such as I2C, SPI, UART, etc., to implement various embedded applications.

5 Results and Discussion

The implementation of the complete house boat built with security system is shown in Fig. 6.

Figure 7a–c depicts the boat structure fixed with sensors to sense in three directions, namely forward, left, and right obstacle. Figure 7d shows the complete setup with sensor locations, including battery, a motor, an indicating device, switch inter connected for operation with the Arduino board. The ultrasonic sensors are fixed such that they detect the obstacle present in front of the boat, towards the left, and towards the right.

The operation of the prototype is tested in a pond for its functionality as shown in Fig. 8. Once the obstacle is detected the boat takes a different safe direction for its movement with an alarm stating the situation of the presence of obstacle in the event of emergency where there is no opportunity for its moment or helpless situation and a signal is given to the control room for assistance.

6 Conclusion and Future Scope

In the present work, a prototype of safe and secure waterways transportation is presented. The developed module consists of a boat with ultrasonic sensor to detect the obstacle in three directions of its moment, i.e. front, left, and right. The existence of the obstacle is detected by the ODM which then gives an RF signal to the control room indicating the emergency. Looking at the notification, necessary action will be initiated by the control room for security purpose. SONAR technique is implemented
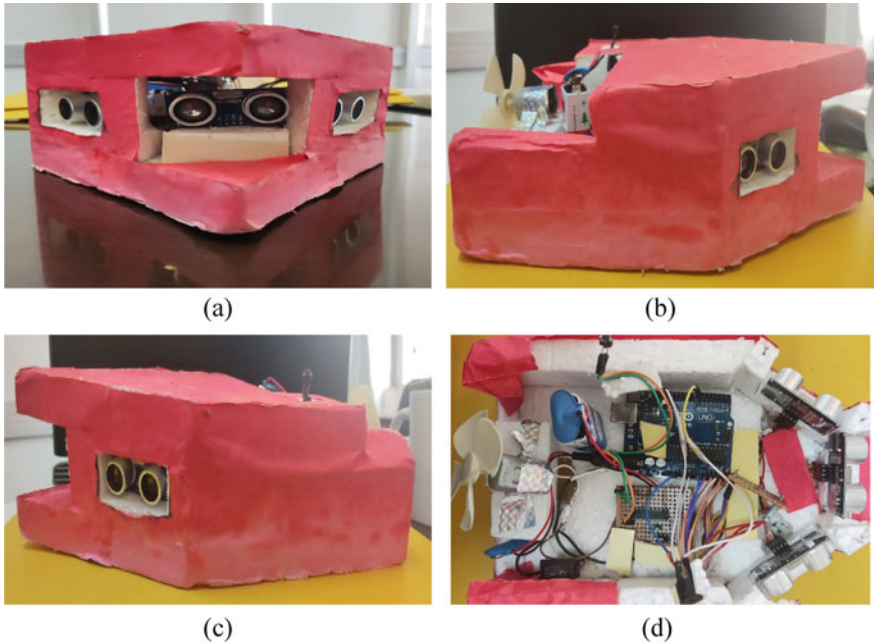
Fig. 7 **a** Front end of the boat depicting the sensors for obstacle detection in three directions, **b** side view of the boat with sensor for sensing obstacle on the right side of the boat, **c** side view of the boat with sensor for sensing obstacle on the left side of the boat, and **d** complete setup with sensors and associated electronic circuitry for boat movement

Fig. 8 Testing the prototype model: **a** boat detecting an obstacle and **b** boat taking a deviation after detection

in the present work to define the size and distance of the obstacle which provides sensing and making diversion from the obstacle. The implemented module can be deployed in passenger vessel boat, cargo vessels, and fish boats for security purpose. As a future scope of our work, the present work can be extended by increasing the span of obstacle location using sophisticated sensors for improved security.

References

1. Ahmed MB, Cavas M (2019) Design and simulation of four walls crossed security system against intrusion using PIC microcontroller. Int J Eng Sci Invention (IJESI) 8(03):59–74
2. Ahmad MB, Bello MI, Muhammad AS, Bello A (2021) Design and simulation of crossed walls security detection system. Int J New Comput Architectures Appl Soc Digital Inf Wireless Commun 11(1):10–21
3. Baballe MA (2021) Detection of crossed walls security alarm system against invasion. Rev Comput Eng Res 8(1):14–26
4. Baballe MA, Cavas M (2018) Design and simulation of four walls crossed security system against intrusion using pic microcontroller. Am J Eng Res 7(12):233–244
5. Singh N, Gunjan VK, Chaudhary G, Kaluri R, Victor N, Lakshmanna K (2022) IoT enabled HELMET to safeguard the health of mine workers. Comput Commun 193:1–9
6. Babelle MA, Bello MI, Akar BA, Sule AT, Mohammed AS (2022) Implementation of security alarm using arduino with P.I.R motion sensor and GSM module. Int J Artif Comput Intell 2(2):1–7
7. Prasad PS, Beena Bethel GN, Singh N, Kumar Gunjan V, Basir S, Miah S (2022) Blockchain-Based Privacy Access Control Mechanism and Collaborative Analysis for Medical Images. Secur Commun Netw, 2022
8. Paulet MV, Salceanu A, Neacsu OM (2016) Ultrasonic Radar. In: International Conference and Exposition on Electrical and Power Engineering (EPE), pp 551–554
9. Lakshmanna K, Shaik F, Gunjan VK, Singh N, Kumar G, Shafi RM (2022) Perimeter degree technique for the reduction of routing congestion during placement in physical design of VLSI circuits. Complexity 2022, 1–11
10. Abdullah RH (2015) Design and implementation of ultrasonic based distance measurement embedded system with temperature compensation. Int J Emerg Sci Eng (IJESE) 3(08):30–37
11. Singh N, Gunjan VK, Nasralla MM (2022) A parametrized comparative analysis of performance between proposed adaptive and personalized tutoring system "seis tutor" with existing online tutoring system. IEEE Access 10:39376–39386
12. Raišutis R, Tumšys O, Kažys R (2010) Feasibility study of application of ultrasonic method for precise measurement of the long distances in air. Ultragarsas J (UJ) 65(1), pp 7–10
13. Sivagnanam G, Midhun AJ, Krishna N, Maria G, Samuel R, Anguraj A (2015) Coast guard alert and rescue system for international maritime line crossing of fisherman. Int J Innovative Res Adv Eng (IJIRAE) 2(2)
14. Jameela T, Athotha K, Singh N, Gunjan VK, Kahali S (2022) Deep learning and transfer learning for malaria detection. Comput Intell Neurosci, 2022
15. Rahman S (2017) An analysis of passenger vessel accidents in bangladesh. In: 10th International Conference on Marine Technology, MARTEC, vol 194, pp 284–290
16. Rahaman NA, Rosli HZ (2014) An innovation approach for improving passengers vessels safety level overload problem. Int J Bus Tourism Appl Sci 2(2)
17. Cheng F, Hirdaris S (2012) Improvement of ship safety through stability research and innovations. In: 11th International Conference on the Stability of Ships and Ocean Vehicles, pp 23–28
18. Ahola M, Romanoff J, Kujala P, Remes H Varsta P (2011) Cruise and Ferry Experience program for future passenger ship designers. In: RINA Conference on Education and Professional Development of Engineers, UK
19. EK A (2003) A study of safety culture in passenger shipping. The 3rd Safety and Reliability International Conference, The publishing and printing house of air force institute of technology 3 pp 99–106

Video Surveillance-Based Underwater Object Detection and Location Tracking Using IoT

Srinivas Samala, V. Ruchitha, D. Sai Pavan, S. Hemanth, V. Soumya, and Syed Musthak Ahmed

Abstract The underwater surveillance gadget displays images of the waterbed. Underwater rovers are being used in a variety of fields, including underwater motion capture systems and aquatic surveillance bots. Underwater object detection still faces issues such as blurred image rendering, texture distortion, and light visibility distortion. Underwater photographs are characterised by poor visibility, low contrast, and distorted information due to severe optical attenuation and light scattering induced by the aqueous medium and suspended particles. The above issues restrict the development of underwater rovers in object detection and transmission of data to the users. The main objective of this paper is to locate and share the visual information of the object underwater. In this project video is continuously monitored and displayed in real-time mode. This live video can be shared to other people where this information can be very useful in emergency situation to help other people. In this system video surveillance and object recognition are made to extract the information about objects underwater in real time. The video surveillance is very helpful to observe all the activities under water and can also be very help for proof of evidence for suspicious activity. This work can be applicable for proper sharing of underwater resources by visual information in real time. It also shares brightness levels under the water. Camera used in prototype provides real-time visuals of underwater ecosystem and in extension to that geo-locations of respective places of water bed will be sent through an SMS to the registered mobile number.

Keywords Underwater rovers · ESP 32 cam · GPS · Object detection

1 Introduction

Underwater robots are being used for a wide range of tasks, like target capture, investigation, and search. In recent years, several technologies have been developed to overcome issues associated with underwater imaging and to undertake autonomous

S. Samala (✉) · V. Ruchitha · D. S. Pavan · S. Hemanth · V. Soumya · S. M. Ahmed
S R Enginnering College, Warangal, India
e-mail: srinu486@gmail.com

assessment of underwater surroundings or underwater vehicle guidance. The major obstacles to underwater perception are the greater costs of the devices, their laborious setup, and the signal and light transmission distortion caused by the water medium. The effects of absorption and scattering on light propagation in underwater environments, in particular, have a profound impact on visual experience. Several approaches for tracking and detecting underwater objects have emerged, and one such approach is sonar, the most prevalent method of localisation and scene reconstruction. A 3D sonar imaging system for object detection is carried out using multi-frequency acoustic wave emissions and a sonar array camera. The following sections present the various works carried out on underwater surveillance using techniques such as ultrasonic underwater technologies and artificial vision.

2 Literature Survey

Underwater images are a valuable tool for hydrographic investigation, as they show water depth and marine possibilities across a large area. Using reference and target photographs, this article offers a versatile technique for detecting a given object inside a cluster of objects. It makes it easy to follow an object's path in a streaming video. This article explains how to recognise a specific object within a group of objects [1]. The referenced article [2] that developed an algorithm is capable of automatically detecting a pipeline on the seafloor as well as some objects in its surroundings, such as bridges and anodes. To address issues related to light attenuation in the water, a colour compensation process has been implemented. Artificial neural networks are then used to classify the pixels in the source images into several classes in real time, each of which corresponds to a different item in the observed scene. Foresti, the referenced article [3], describes how it is challenging to detect and recognise underwater objects using computer vision in low-light environments. Intervention missions that call for the grasping and manipulation of underwater objects necessitate an AUV vision system that is capable of object identification, localisation, and tracking. Assembling the MARIS AUV's vision system and using it to identify common cylindrical pipes.

The referenced article [4] suggested a method for analysing acoustic and visual data in order to obtain information about the presence of constructed and archaeological artefacts on the seafloor. This method statistically highlights such distortions in the surrounding environment, appropriately valuing the persistence of significant curves in a video clip or a sonogram. To do this, we used the ELSD algorithm, a parameter-free method based on Gestalt principles that have shown to be effective. The referenced article [5] demonstrated the utility of vision-based object detection techniques in aquatic images using a variety of datasets, as well as the limitations that arise in diverse scenarios. Light transmission in water creates distortion and attenuation, as well as difficult working circumstances for underwater computer vision. Finding a target object using this method relies on properties that are unaffected by the underwater environment, such as strong colour consistency and distinct edge

definition. We test the performance of the proposed approach on a variety of underwater datasets. Different stereo camera quality levels were used to build datasets with varying target item types and colours, as well as acquisition depth and circumstances [6].

Using forward-looking imaging sonar, the referenced article [7] proposed a real-time submerged object detecting technique. Based on the features of sonar images, the Haar-like feature is used to construct each weak classifier [8]. The adaptive boosting (AdaBoost) technique is used to calculate the parameters of every base classifiers and the scores of the training sample. The referenced article [9] applied deep machine learning and has found lot of success in object detection and categorisation in digital images. The characteristics and machine learning architectures utilised are emphasised, and the analytical methodologies are classified per the object of detection [10].

The referenced article [11] used an integral-image representation strategy to quickly compute parameters and reduce computing burden as the detection process progressed, utilising the proposed approach. It is possible to detect sonar data in real-time onboard vehicle. The proposed method was built and integrated into the Gemellina remote-controlled vehicle's software system to run in real time [12].

The referenced article [13] made use of geophysical direct-current impedance techniques to develop a new real-time underwater item detecting method. Their ultimate goal is to identify and track tiny submarines in real time in sea to a depths of 100 m and less using acoustic signalling [14]. The following are the primary characteristics of our method: fixed interelectrode setups enabling high-speed real-time object recognition and tracking of extremely low-level electric field signals are their main features. The referenced article [15] used coastal defence which is quite interesting in undersea object identification and tracking. In audibly loud situations a new mechanism namely direct-current resistivity survey methods has been implemented in place of traditional detection approach. The referenced article [16] research suggests employing generative adversarial networks (GANs) to enhance the quality of visible aquatic sceneries, with the purpose of improving input to eyesight behaviours farther down the autonomous pipeline. Their methods are used to create a dataset for underwater image enhancement. The referenced article [17] presented a flexible detection technique to enhance the mean average precision percentage. Their strategy included two fundamental ideas 1 for the performance improvement and other for performance gain. The referenced article [18] proposed YOLO and object detection technique. This novel method surpasses other detection techniques like DPM and R-CNN. The referenced article [19] proposed a genuine method for the detection of images using RGB. In their research they made use of convolutional neural network (CNN) to obtain image accuracy.

The referenced article [20] proposed 2D side-scan sonar imaging that offers a method for classifying and locating marine mines. The technique closely resembles Turk and Pentland's eigenfaces technique, and it approaches mine detection as a two-dimensional object recognition and localisation problem, recognising that mine patches have a degree of uniformity in size and texture. In [21–26], several other works on underwater object detection and capturing images are presented.

Fig. 1 Block diagram of the
proposed method
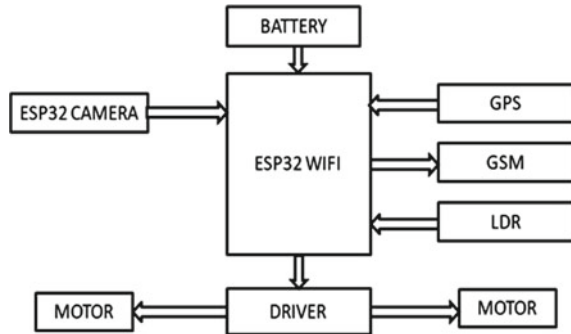
3 Proposed System

This project is to design an underwater surveillance device that provides visuals of a
water bed. It shares the visuals of respective places that could be used to detect unusual
and suspicious movements. It also shares brightness levels under the water. Here we
use two devices, one device to monitor the sea bed and another device to control
the movement. Camera provides real-time visuals of the underwater ecosystem. It
also provides geo-locations of respective places through an SMS to the user. This
proposed system is made of using ESP32CAM, GPS module, GSM module, LDR,
L293 motor driver, IC and DC motors. This device will also provide light condition
levels inside the water bed and also provide an optimised light. Figure 1 depicts the
proposed system in a block diagram.

4 Results and Discussion

Figure 2 shows details of the prototype devices in developing the module. The module
is divided into four sections, i.e., control section, propallent section, communication
section, and submerging section.

Control section: The control section consists of microcontroller which controls the
sensor activities and propellant control. The microcontroller utilised is ESP 32. The
ESP32CAM module is developed to check an underwater object using IoT. The
location of the object is tracked using wireless communication technology (GSM
and GPS). This is shown in Fig. 3.

Communication section: The communication section consists of GSM module
which communicates with the user by sending the data through an SMS to the regis-
tered sim. The operational voltage of the communication is 12 V DC supply. This is
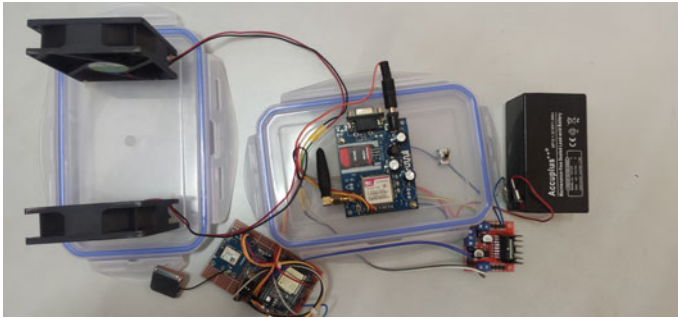shown in Fig. 4.

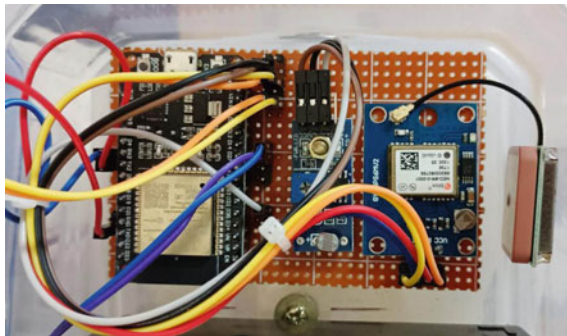**Fig. 2** Discrete parts of the prototype of underwater object detection and location tracking system
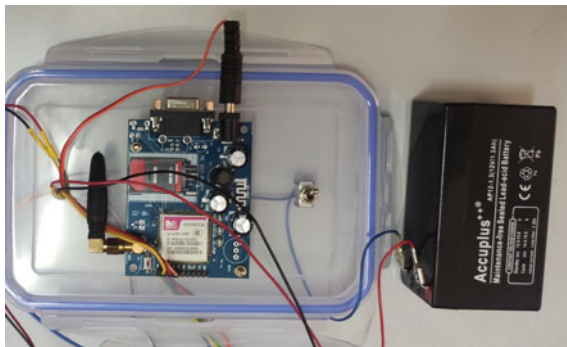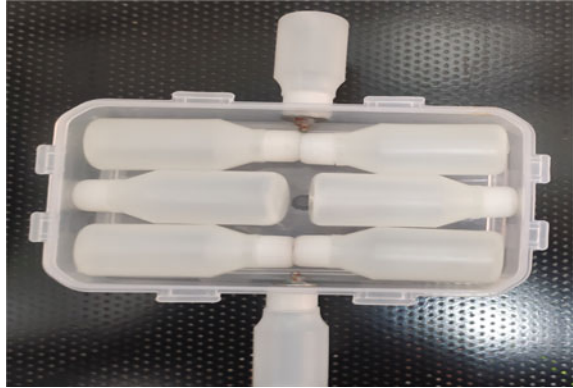

Fig. 3 Control section of the prototype

Fig. 4 Communication section of the prototype

Submerging section: The submerging section consists of air sealed bottles inside. This section will submerge the device partially into the water. The submerging section

Fig. 5 Submerging section of the prototype

is sandwiched between propeller section at the bottom and the control circuit placed at the top. This is shown in Fig. 5.

Propelling section: The propeller section is fixed at the bottom of the boat module. It consists of two propellers which will make the device to move forward and backward directions, controlled by the driver circuit as shown in Fig. 6.

The complete setup of Video Surveillance-Based Underwater Object Detection And Location Tracking System is shown in Fig. 7. The testability is carried out in a pool as shown in Fig. 8. The presence of obstacle is detected by the visual device fixed underneath the submerging section. The results of visuals are shown in Fig. 9 and 10.

The device is first turned on by a switch button. Beforehand the user must download the appropriate application to their mobile device and pair it with the device using the IP provided by the device. Once connected, the application's buttons will function to operate the device. Once the obstacle/object is detected the information is sent to the user through Wi-Fi streaming. The user can get the location details
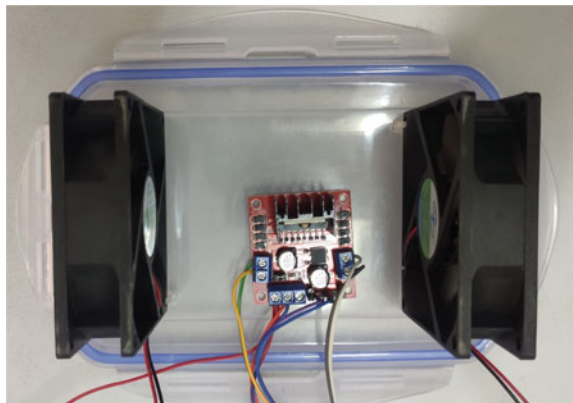
Fig. 6 Propeller section of the prototype

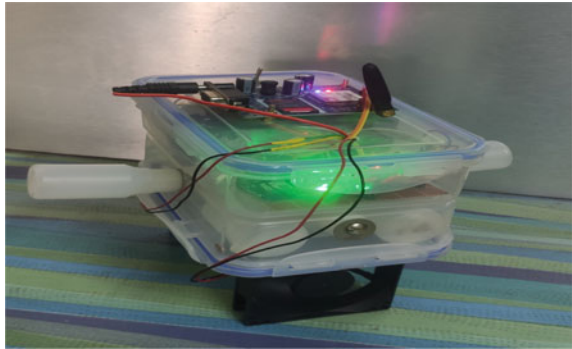**Fig. 7** Underwater object detection using IoT

Fig. 8 Controlling the movements of device using Blynk App

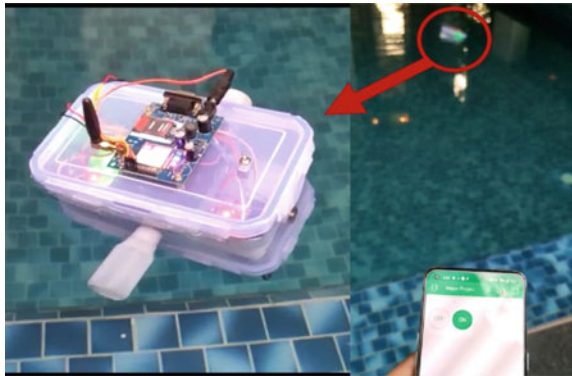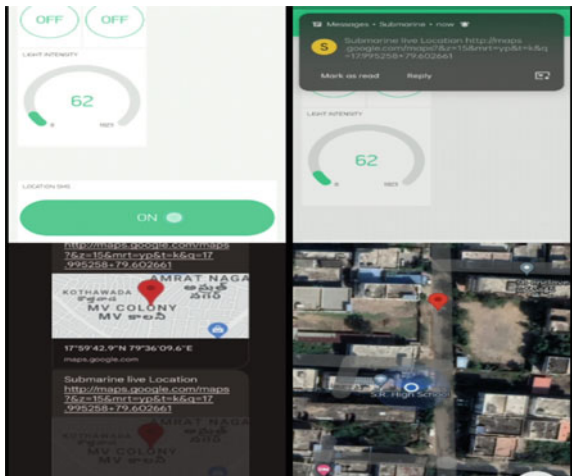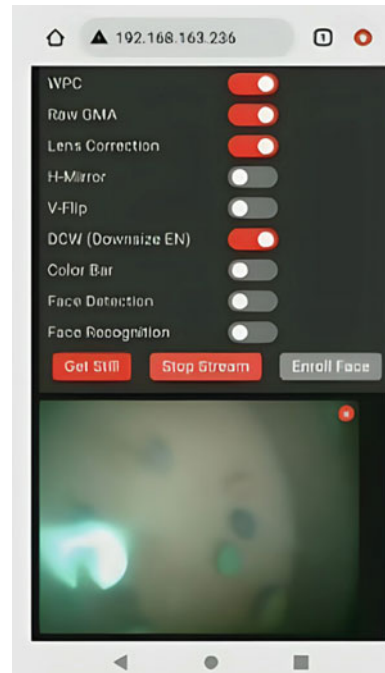**Fig. 9** Device live location sharing via SMS

Fig. 10 Live underwater
video streaming over LAN
through Blynk App

through an SMS to the mobile. The prototype communicates with open-source cloud platform (Blynk) and the corresponding mobile App. Blynk is used to control the moment of the prototype. The forward and backward buttons are used to control the device's motion, and a location tab button is provided to send the device's location and is displayed on the mobile app.

5 Conclusion

In the present work, an underwater surveillance device to capturing objects with improved clarity, texture, and clear visibility is presented. The objective is to provide an accurate data of the suspicious objects under water. The implemented prototype shares the visuals of respective places that could be used to detect unusual and suspicious movements. The camera used in prototype provides real-time visuals of underwater ecosystem and in extension to that geo-locations of respective places of water bed. The objective to locate and share the visual information of the object underwater by continuously monitoring and displaying in real-time mode is achieved. The live video provides valuable information that can be very useful in emergency situation to help other people. This video surveillance helps to study activities under water by providing necessary information about suspicious activity under water.

This work can be applicable for proper sharing of underwater resources by visual information in real time.

References

1. Rizzini DL, Kallasi F, Oleari F, Caselli S (2015) Investigation of vision-based underwater object detection with multiple datasets. Int J Adv Rob Syst 12(6):77
2. Li M, Ji H, Wang X, Gong Z (2013) Underwater object detection and tracking based on multi-beam sonar image processing. In: 2013 IEEE international conference on robotics and biomimetics (ROBIO). IEEE, pp 1071–1076
3. Foresti GL, Gentili S (2000) A vision based system for object detection in underwater images. Int J Pattern Recognit Artif Intell 14(02):167–188
4. Skaldebø M, Haugaløkken, BOA, Schjølberg I (2019) Dynamic positioning of an underwater vehicle using monocular vision-based object detection with machine learning. In: OCEANS 2019 MTS/IEEE SEATTLE. IEEE, pp 1–9
5. Singh N, Gunjan VK, Chaudhary G, Kaluri R, Victor N, Lakshmanna K (2022) IoT enabled HELMET to safeguard the health of mine workers. Comput Commun 193:1–9
6. Mahavarkar A, Kadwadkar R, Maurya S, Raveendran S (2020) Underwater object detection using tensorflow. In: ITM web of conferences. EDP Science, vol. 32, p. 03037
7. Kim B, Yu SC (2017) Imaging sonar based real-time underwater object detection utilizing AdaBoost method. In: 2017 IEEE underwater technology (UT). IEEE, pp 1–5
8. Prasad PS, Beena Bethel GN, Singh N, Kumar Gunjan V, Basir S, Miah S (2022) Blockchain-based privacy access control mechanism and collaborative analysis for medical images. Secur Commun Netw
9. Moniruzzaman M, Islam SMS, Bennamoun M, Lavery P (2017) Deep learning on underwater marine object detection: a survey. In: International conference on advanced concepts for intelligent vision systems. Springer, Cham, pp. 150–160
10. Galceran E, Djapic V, Carreras M, Williams DP (2012) A real-time underwater object detection algorithm for multi-beam forward looking sonar. IFAC Proc Volumes 45(5):306–311
11. Jameela T, Athotha K, Singh N, Gunjan VK, Kahali S (2022) Deep learning and transfer learning for malaria detection. Comput Intell Neurosci, 2022
12. Cho SH, Jung HK, Lee H, Rim H, Lee SK (2016) Real-time underwater object detection based on DC resistivity method. IEEE Trans Geosci Remote Sens 54(11):6833–6842
13. Lee H, Jung HK, Cho SH, Kim Y, Rim H, Lee SK (2018) Real-time localization for underwater moving object using precalculated DC electric field template. IEEE Trans Geosci Remote Sens 56(10):5813–5823
14. Lakshmanna K, Shaik F, Gunjan VK, Singh N, Kumar G, Shafi RM (2022) Perimeter degree technique for the reduction of routing congestion during placement in physical design of VLSI circuits. Complexity, 2022, 1–11
15. Fabbri C, Islam MJ, Sattar J (2018) Enhancing underwater imagery using generative adversarial networks. In: ICRA
16. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. Extended version of our CVPR 2014 paper.
17. Singh N, Gunjan VK, Kadiyala R, Xin Q, Gadekallu TR (2022) Performance evaluation of SeisTutor using cognitive intelligence-based "kirkpatrick model". Comput Intell Neurosci, 2022
18. Redmon J, Divvala S, Girshick R et al (2016) You only look once: unified, real-time object detection. In: IEEE conference on computer vision and pattern recognition. IEEE Computer Society, pp 779–788
19. Zhou X, Lan X, Zhang H, Tian Z, Zhang Y, Zheng N (2018) Fully convolutional grasp detection network with oriented anchor box. arXiv:1803.02209

20. Singh N, Gunjan VK, Nasralla MM (2022) A parametrized comparative analysis of performance between proposed adaptive and personalized tutoring system "seis tutor" with existing online tutoring system. IEEE Access 10:39376–39386
21. Saisan P, Kadambe S (2008) Shape normalized subspace analysis for underwater mine detection. In: Proceeding of IEEE ICIP 2008, vol 1, pp 1892–1895
22. Zhang T, Liu S, He X, Huang H, Hao K (2019) Underwater Target Tracking Using Forward-looking Sonar for Autonomous Underwater vehicle (AUV). Nat Key Lab Sci Technol 1:1–28
23. YaŞar FG, Kuseto Ğullari H (2018) Underwater human body detection computer. In: 26th Signal processing and communications applications conference (SIU), pp 1–20
24. Gaude G, Borkar S (2018) Comprehensive survey on underwater object detection and tracking. Int J Comput Sci Eng 6:304–313
25. Chen Z, Zhang Z, Dhai F, Bu Y, Wang H (2017) Monocular vision-based underwater object detection. Int J Comput Sci Eng (JCSE), pp 1–5
26. Foresti GL, Gentili S (2016) A vision based system for object detection in under water images. Int J Pattern Recognit Artif Intell 2:167–188

Pothole Detection and Warning System for Intelligent Vehicles

Jatin Giri, Rohit Singh Bisht, Kashish Yadav, Navdeep Bhatnagar, and Suchi Johari

Abstract Due to inadequate road maintenance, the road conditions are terrible everywhere. This is particularly true for the urban/rural roads. The potholes on the road grow bigger and deeper with each monsoon, increasing the number of road accidents. Over the last few years, many researchers have proposed numerous solutions to this problem. Researchers have developed systems to detect potholes in real time. This data is recorded and passed to the road authorities by some of the researchers. The recorded data, if stored in the cloud, requires network connectivity. Also, this data is not used for reducing the number of accidents due to potholes. In this paper, we have proposed a strategy where information on the road stays on the road and will be used to notify the vehicles about road conditions to avoid accidents. We are emphasizing on the importance of storing data locally rather than on remote servers. Real-time scenarios have been identified and simulated for the study. The results obtained after simulation show that the number of road accidents caused by potholes can be successfully reduced.

Keywords Potholes · VANET · Vehicular communication · Road accidents · Accelerometer · Automated vehicles

J. Giri · R. S. Bisht · K. Yadav
Department of Computer Science and Engineering, Tula's Institute, Dehradun, India

N. Bhatnagar
School of Business, University of Petroleum and Energy Studies, Dehradun, India
e-mail: navdeepbhatnagar84@gmail.com

S. Johari (✉)
School of Computing, DIT University, Dehradun, India
e-mail: shuchi.johari@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
A. Kumar et al. (eds.), *Advances in Cognitive Science and Communications*,
Cognitive Science and Technology, https://doi.org/10.1007/978-981-19-8086-2_113

1   Introduction

There are different modes of transportation available around the world, but road transportation is the most popular. Roads play an essential role in economic development and provide important social advantages such as conveying goods and passengers, as well as contributing to Gross Value Addition (GVA). Road transportation accounts for 3.1% of the total contribution to Gross Value Addition (GVA) from the entire transportation sector, which totals 4.63%. With a length of around 63.86 lakh km, India's road network is the world's second-largest network, comprising National Highways, Expressways, State Highways, District Roads, Rural Roads, Urban Roads, and Project Roads [1]. Since the use of motor vehicles is rapidly increasing, road safety is deteriorating at an equal pace [2] and has become a concern. The government, particularly the motor industry, is working to improve passenger safety and reduce road accidents. Over-speeding, drunken driving, distractions from fellow passengers, rash or reckless driving, and failure to wear safety equipment such as helmets and seat belts are all factors that contribute to accidents [3]. According to the National Crime Records Bureau (NCRB), approximately 354,796 road accidents occurred in India in 2020, resulting in 133,201 deaths and 335,201 injuries. Over-speeding was the cause of more than 60% of those accidents, resulting in 75,333 deaths and 209,736 injuries, and 24.3% of road accidents were caused by dangerous and careless driving, resulting in 35,219 deaths and injuries. Figure 1 depicts and illustrates the year-wise accidents, deaths, and injuries on the road.
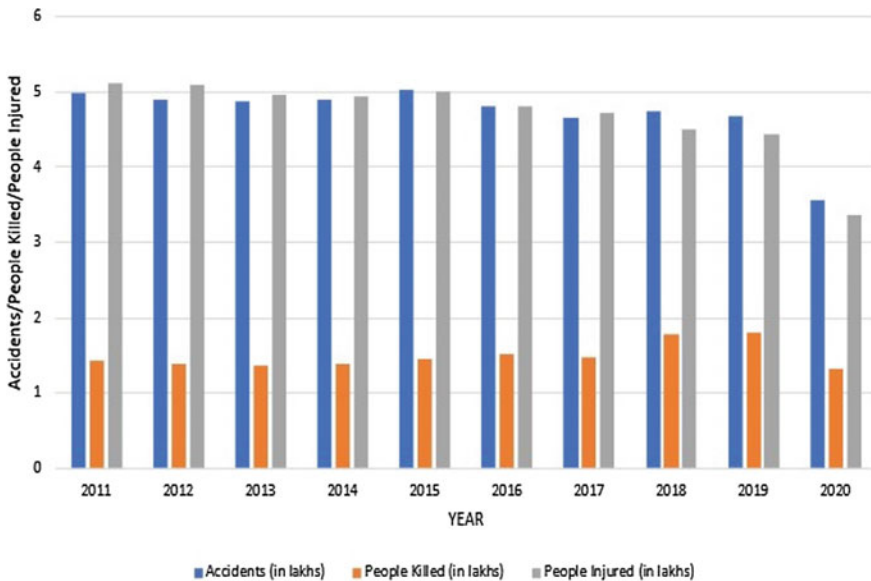

Fig. 1 Road accidents (2011–2020)

In comparison to India (36 lakh), developed countries such as the United States (684 lakh), Japan (393 lakh), and Turkey (233 lakh) have a far larger number of accidents [4]. Even with COVID-19 causing lockdown, this is a big number. According to UN Economic and Social Commission for Asia and the Pacific (UNESCAP) reports, India loses about 3% of its GDP each year due to road accidents, which amounts to about $58 billion in value terms. In terms of GDP loss, India is only second to Japan, which loses about $63,000 million, and only Iran is ahead of India in terms of GDP loss, which is about $30,697 in value terms [5]. This paper is divided into seven sections. Section II discusses about the existing work done by various researchers in different fields of VANETs, etc. Section III presents the proposed approach of the paper. Section IV discusses the simulation, and Section V shows the results, whereas Section VI covers the problems and future work. The conclusion is presented in Section VII. Most road accidents are caused by driver's recklessness, although it is not always their fault, accidents can also be caused by broken roads or potholes [6]. A pothole is a section of a road that has evolved into a hole because of a fissure in the road surface being worn away [7]. Water trapped on roads and overloaded vehicles are the major causes of potholes, which contribute to surface degradation and erosion of rocks beneath the road surface [8], and tiny cracks in the road are not repaired right away. The size of the potholes varies depending on where you are; some are minor, while others are rather large, but all of them are incredibly dangerous. Potholes may be seen on almost every road in India; hence, the term is not unfamiliar to Indians. We are all aware of the problem, and how it is magnified by poor road construction in adverse weather. It gets worse during the monsoon. The roads were waterlogged. We could not even see the road back then, so how could anyone spot these potholes? According to the Ministry of Road Transport and Highways, over 9300 persons have perished and over 25,000 have been injured because of these hidden monsoon potholes in the last three years [9]. Apart from that, potholes create a lot of damage to our vehicles as well, such as tire D-shaping, which makes the tire weak and requires replacement as soon as possible, otherwise and the weak tire can burst. Tire replacement is not something that everyone can afford regularly. Potholes also impair the car's suspension, bolts, and shocks, requiring them to be replaced considerably sooner than anticipated [10]. We can avoid some of these potholes by using proper handling tactics, but we will not be able to prevent all of them because most potholes are discovered at the last possible moment. We cannot blame the government, authorities, or anyone else for the threats we and our vehicles face [11]. In this current era, where technology has unquestionably benefited us greatly, a pothole detecting system is essential to alert drivers about potholes [12]. Detecting potholes in a precise, fast, and efficient manner is a critical task [13]. The purpose is to prevent pothole-related road accidents, save cars from damage, and assist drivers in anticipating potholes [14].

A. VANET

VANET, or Vehicular Ad hoc Network, is a type of Mobile Ad hoc Network (MANET) that offers hope for the future of intelligent transportation systems through inter-vehicle communication [15]. As an emerging field, VANET is becoming a

Fig. 2 VANET and its components
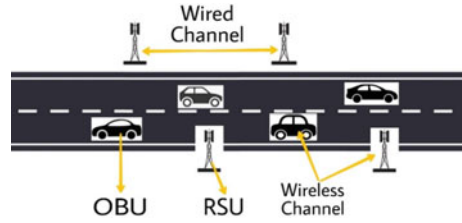
crucial component of intelligent transportation systems [16]. The use of MANET routing protocols to adapt VANET has become common [17]. Flexibility, scalability, and multitenancy are three fundamental qualities of a VANET that aid in its efficient deployment [18]. Because VANET incorporates wireless communication between cars, information security is a critical concern [19]. Government agencies have designated certain frequency bands for vehicular communication. Vehicles communicate with one another and share information [20]. Different types of communication in VANET include vehicle-to-vehicle (V2V), vehicle-to-RSU (V2R), vehicle-to-infrastructure (V2I), etc. communications. Routing has a significant role in the performance of VANET networks [21]. Each VANET node acts as a router to facilitate multi-hop communication among other nodes. Because there are no central access points, communication between nodes can be direct (single-hop) or through other nodes (multi-hop). VANET safety applications are provided by IEEE 802.11 specialized Short-Range Communication, Medium Access Control Layer, and Physical Layer standards [22]. VANETs provide short to medium-range communication systems [23]. The OBU and RSU nodes exchange information over a wireless channel as shown in Fig. 2.

B. OBU

The acronym OBU stands for on-board unit. These are the VANET units that are in motion. OBUs are low-power devices that employ Dedicated Short-Range Communication Services (DSRCS) to send data to roadside units to improve traffic flow and safety, as well as for other intelligent transportation system goals. The DSRCS on-board units are licensed to operate in the 5895–5925 MHz band [24]. It is a hardware component found in intelligent vehicles. The primary function of these units is to communicate with other OBUs and RSUs.

III. RSU

The acronym RSU stands for roadside unit. RSUs help in communication and are placed at a set distance on both sides of the road [25]. RSUs have an antenna, processor, and read/write memory, as well as wired and wireless interfaces. Wireless interfaces are used for communication of the RSU with OBU [20]. The wired interface is used to connect to other RSUs and the internet [20]. RSUs can be easily exploited because they are installed along the roadway, making them untrustworthy [26].

2 Related Work

Researchers have been studying the topic of road defects for several years. Their goal is to ensure that these potholes do not endanger the driver's life or the vehicle itself. Many solutions surfaced, and we divided them into two categories. The first set of technologies targeted at detecting these flaws and relaying that information to the driver or automated vehicle in real time. These will aid in the execution of important changes, while reducing negative consequences. All of this occurred in real time, while the vehicle was on the road. Vehicles are equipped with sensors that transmit road information to the host vehicle. People have employed a variety of sensors, like ultrasonic, accelerometer, gyroscope, and LiDAR. The system will sense its surroundings using the sensors, and then all that is left is a decision-making procedure based on the information it has gathered. The information that has been recorded is processed for this. This has been accomplished through a variety of methods:

- **Machine Learning:** The first way is based on machine learning, which is a relatively new field. Researchers have developed approached based on unsupervised algorithm or data processing [27]. The machine can recognize patterns in the data it collects and categorizes which parts of the road are potholes and which are not. Then, there are supervised algorithms, which require the system to be trained first using already categorized data to improve detection accuracy for road defects. Researchers presented a random forest approach in [28], whereas they used neural networks in [29, 30]
- **Image Processing:** The second method is based on the image processing. Because noise in collected photographs is unavoidable, image processing procedures are required to filter the data. Threshold segmentation is one of the image processing algorithms used under this approach [31].
- **3D Processing:** Third method is based on the 3D data. This approach works by recognizing 3D objects in space and then detecting potholes from that. LiDAR, which stands for Light Detection and Ranging, is the sensor utilized in these. One of the strategies is shown in the [32] study, which is applied with convolutional neural networks. This newly created data will assist in real time and can be recorded so that other vehicles can use it to avoid potholes. The researchers employed a combination of 2D picture and 3D surface modeling approaches in [33].

In paper [34], an excellent overview of these and other strategies is discussed. The second approach does things a little differently. Instead of providing real-time security, these are based on storing defect data. Though not as widespread as the first, these solutions try to reduce the likelihood of road accidents by providing defect data to the vehicle [35] or driver ahead of time. The information is also given to government officials so that it can be used for road maintenance. The authors of the papers [35–37] developed a technique for storing spatial information about potholes using GPS and then utilizing that information afterward. These values are

saved to the cloud so that they can be utilized by the government or to assist drivers through an app (we will see how our simulation compares to their work later, but it depicts different things), whereas the others have just detected the potholes and their depth. They have nothing that can pinpoint the exact location of potholes on the road. Many studies have employed mobile phones and their built-in sensors to detect these problems, as shown in [8, 35–43]. Potholes are detected by these sensors only after the car has hit them, not before. However, these sacrifices will be required to obtain road information that may be used for road maintenance or kept so that consumers can access it via an app. The authors of [44–48] also conducted study on over 150 potholes and over 100 speedbumps, categorizing them according to their safety levels. This is also something that might be used in the future studies where we could quickly identify a pothole and its level of caution in real time. Regression and classification are examples of supervised learning techniques employed by the authors [49–52]. This method makes use of mobile devices, and the data collected has been used to categorize the data using supervised learning algorithms, as we have seen earlier. This information is utilized to train the algorithm to detect potholes and determine their depth. Assist in improving the efficiency of data detection methods for real-time detection of potholes. The different protocols followed in VANETs are discussed in [53–55].

3   Proposed Approach

The proposed approach belongs to the second class of algorithms discussed in the related work. Proposed approach is differentiated from other approaches as data is not stored on the server. First and foremost, we must state that our research is not focused on how we read data regarding potholes, but rather on how we store and utilize that data for greater efficiency. The data collection procedure is totally independent, making it extremely efficient. Our main idea is to store the data locally! To understand this, refer Fig. 3; let us say there is a road segment AB with RSUs at each end. So, we have one at point A and another at point B. The proposed approach is divided into four phases: (1) pothole detection phase, (2) pothole warning phase, (3) vehicle action phase, and (4) pothole update phase. Both these phases are independent.

A.   Pothole Detection Phase

The main purpose here is to have a pothole detection and data storage mechanism to store the spatial location of the detected pothole that must be available with the
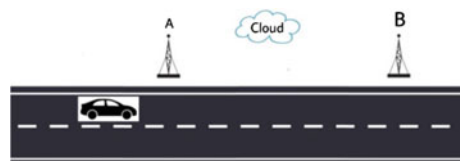
Fig. 3 Road segment with RSU at its end
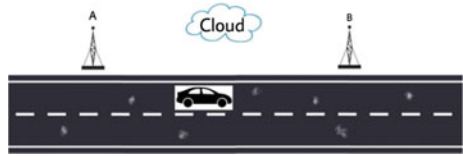
Fig. 4 Car detecting potholes

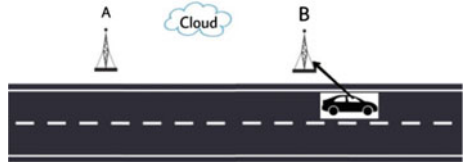**Fig. 5** Vehicle passing spatial information to the RSU at B

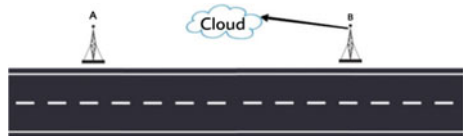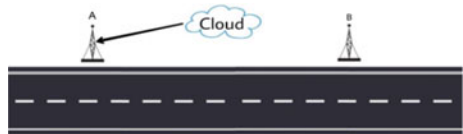**Fig. 6** RSU storing data in the cloud


Fig. 7 RSU retrieving information

vehicle that passes this road segment as shown in Fig. 4. For any stretch, when the vehicle enters the network, it registers itself to the RSU. For any pothole detected by the vehicle for the first time, the data will be captures. The details of the pothole detected are send to the registered RSU. As shown in Fig. 5 for the road segment AB, the details of the potholes are sent to the RSU B as the vehicle is currently registered to RSU B.

Similarly, a car going from B to A stores the collected information at the RSU A. The capabilities of RSUs are increased with the amount of memory to store the captured data. For now, let us assume that the collected information on the potholes is passed from the RSU to the cloud as shown in Fig. 6.

B. Pothole Warning Phase

In the pothole warning phase, the data collected for the potholes is used to warn the vehicles of the existing potholes so that immediate action can be taken and accidents can be avoided. When the vehicle enters the road segment, the RSU at *A* will send the spatial information of the defects on the road to the vehicle. RSU first reads the data from the cloud as shown in Fig. 7, and then forward that information to the vehicle as shown in Fig. 8.

Fig. 8 RSU at the entry point relaying spatial information to a new vehicle
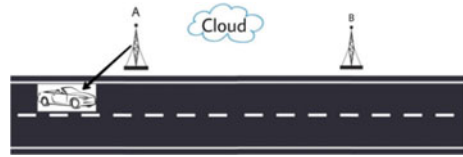
III.   Vehicle Action Phase

The collected information can be used differently in autonomous vehicles and driver-controlled vehicles.

(1) **Autonomous Vehicles**: Once the OBU of the vehicle is notified about the pothole, it checks the details of the detected pothole and how deep it is. Based on the depth of the pothole and the vehicle's capabilities such as shock absorbers, the OBU calculates the speed reduction. The accelerometer reduces the speed of the vehicle as per the instructions of the OBU. Once the pothole is crossed, the OBU will notify the accelerometer to increase the speed again accordingly.

(2) **Driver-Controlled Vehicles**: In driver-controlled vehicles, the details of the potholes will be notified to the driver either on his smartphone or on the vehicle's inbuilt infotainment system. Now the driver is pre-informed of the upcoming pothole, so he has plenty of time to make a decision and reduce the speed accordingly.

IV.   Pothole Update Phase

It might happen that after some time, the pothole is repaired or a new pothole has appeared on the road segment. So, it is equally important to update the stored information. Details of the potholes are updated as:

(1) **For a repaired pothole**: When the OBU is notified about the pothole, it will reduce the speed and follow the instructions, but if it is identified by the vehicle that the pothole no longer exists and there was no need to reduce speed, then it will notify the registered RSU about the same. RSU will update the details accordingly.

(2) **For a new pothole**: Suppose there is a new pothole that was not updated on the RSU. The RSU will not notify the OBU about the same as a result vehicle will continue to move at the same speed that will give a jerk to the vehicle notifying the presence of the pothole. As the OBU knows that the information about this pothole is not there on the RSU, so OBU will send the details to the RSU for the update. Once the details are updated, the further vehicles will be notified accordingly.

This approach can be used for any road segment to notify the vehicles about the defects in advance. The proposed approach reduces the number of accidents caused by potholes. With a vehicle passing by on the road, we can expect a more precise outcome than before. In the proposed approach, we motivate to store the collected information on the RSUs rather than servers. Storing and retrieving data on the server

require internet connectivity. It might face an issue in bad weather conditions, slow internet speed, signal issues, etc. both at the RSU end and server end. So storing data locally at the RSU will resolve the connectivity issues. Vehicles are registered to the RSU, and create a local network without any requirement of an internet connection. This feature makes the proposed approach more reliable than the existing pothole detection approaches where the data is stored over the server.

4 Simulation

For simulation purposes, we have used the accelerometer embedded in smartphones. These smartphones act as the OBU unit for the detection of the potholes in the first phase and display the warning of the detected potholes in the second phase. An Android-based application is created to fetch and store the data of the detected potholes. For the simulation purpose, the local Android Room database of the smartphone is considered as a storage unit RSU. As previously said, the detection technique is independent of our strategy; one can use any detection method. The proposed approach is utilizing the most basic of tactics here, relying on the phone's accelerometer for detection. When the app detects a change in the Z acceleration, it is identified as a pothole. If the acceleration change is greater than 5 but less than 10, the app will notify us that a "Pothole is detected", as shown in Fig. 9. If it is greater than 10, the app will display a notice saying, "Big Pothole is detected", as shown in Fig. 10. The data collected for the potholes is stored in the database. The position of the pothole is identified and stored along with its depth. For simulation purposes, we have considered a road segment with lots of potholes. A vehicle with a smartphone inside it drives from one end of the road segment to another. The smartphone captures the details of the potholes encountered in its local storage. Next time when the vehicle starts from the same point, then the notification is received on the smartphone displaying the detection of the pothole. Along with this, the depth of the pothole and the amount of speed reduction are also suggested to the driver based on his current speed. The log file of the data stored is shown in Fig. 11. Whenever the vehicle continues down the same path, this information is collected from the local database of the vehicle. Based on this information, the application gets an alert in advance for the upcoming pothole as shown in Fig. 12. The speed of the vehicle can be reduced accordingly to avoid any kind of accident.

5 Results

This section shows the comparison result of the proposed approach with three other approaches: CRATER [36], Aid Drivers [37], and Vehicle with Sensors [38]. All these methods used an application that identifies road defects, stores their data, and subsequently warns users about potholes. These approaches are compared based

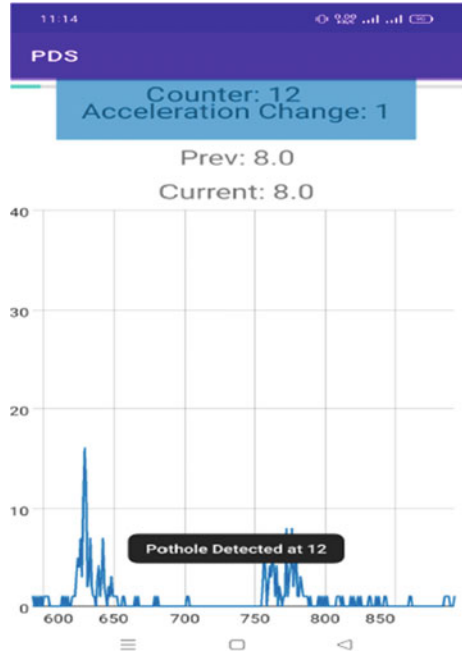**Fig. 9** App displaying
"Pothole Detected" message

Fig. 10 App showing "Big
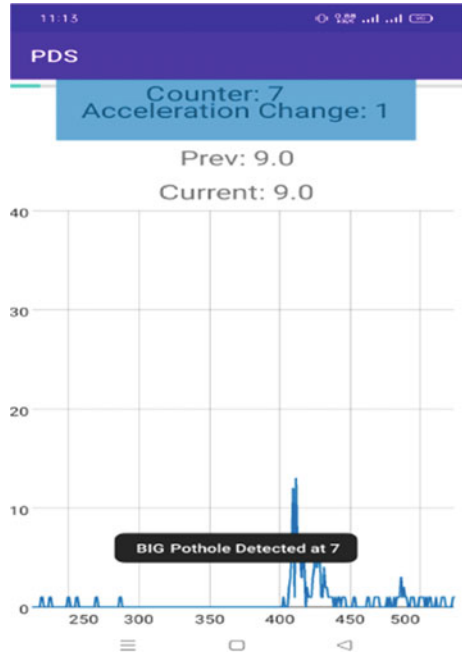Pothole Detected" message

Fig. 11 Log files of all the detected potholes during the first trip

Fig. 12 Alert message for the drivers

Table 1 Assumed values for delays per packet

Parameters	Assumed value
Size of the packet (Mb)	1
Transmission delay (milliseconds)	10
Processing delay (milliseconds)	0.001
Queuing delay (milliseconds)	0.001
Traffic (milliseconds)	0.001

on two parameters, latency and throughput. Latency is defined as the total time taken to detect the pothole, save its information, and send a warning signal to a new approaching vehicle. Latency is calculated as:

$$\text{Latency} = Q_d + 2 * T_d + 2 * P_{\text{Pd}} + \text{Traffic} \tag{1}$$

where Q_d represents queuing delay, T_d represents transmission delay, and P_{Pd} represents propagation delay. Because data is stored and then subsequently retrieved, both propagation and transmission times are doubled. The first approach, CRATER, has an additional processing delay that adds to its latency. It makes use of machine learning and had to calibrate locations using GPS. Similarly the second approach, Aid Drivers has an additional processing time for GPS calibration. The third approach, Vehicle with Sensors, has increased latency due to the extra processing time taken by GPS and ultrasonic sensors. Throughput is inversely proportional to the latency. It is the amount of processing done by the vehicle per unit of time.

$$\text{Throughput} \propto 1/\text{Latency} \tag{2}$$

Table 1 shows the assumption of the values used for the calculation of the approximate values for the delays that are shown in Table 2.

Table 2 Delays values for all the approaches for a single vehicle

Delays	Crater	Aid drivers	Vehicle with sensors	Proposed approach
Number of packets	1	1	1	1
Processing delay (total)	0.002	0.001	0.002	0.001
Queuing delay	0	0	0	0
Transmission delay	20	20	20	20
Propagation delay	0.3336	0.3336	0.3336	0.00003336
Traffic	0	0	0	0
Latency	20.3356	20.3346	20.3356	20.00103336
Throughput	0.049174846	0.049177264	0.049174846	0.049997417

Two scenarios are considered for checking the scalability of these approaches:

A. **Increase in number of vehicles on a single road**: In this, a single road segment is considered and a variable number of vehicles are considered for this road segment. With the increase in the number of vehicles, the number of packets, processing delays, queuing delays, and traffic rise linearly. So the number of vehicles is directly proportional to the mentioned factors. Delay in storing and fetching pothole-related information is directly proportional to the number of vehicles. The calculated delay for storing and fetching information for different approaches for a variable number of vehicles is shown in Table 3. The graphs for latency and throughput for the variable number of vehicles on a single road segment are shown in Figs. 13 and 14, respectively. The latency of the proposed approach is minimum as compared to other approaches, and the calculations have proved that the proposed approach has better throughput. The other three approaches on the other hand have higher latency and lower throughput, because of the additional delays.

B. **Increased number of roads with a constant number of vehicles**: Here, we have considered a variable number of roads, and the number of vehicles per road remains constant. It is calculated that latency increases linearly with the increase in the number of roads for the existing approaches. Existing approaches have a single server for storing and retrieving data for all the road segments; hence, the delay increases linearly. But for the proposed approach, latency does not increase with the number of roads. In the proposed approach, the data is stored and retrieved from the RSU installed independently on every road segment. This lessens the delays involved in fetching the data from the single-loaded server. The delay involved in storing and fetching data of the potholes for variable road segments for these approaches is shown in Table 4. Latency and throughput for a variable number of roads with a constant number of vehicles are shown graphically in Figs. 15 and 16. It is observed through simulation that the proposed

Table 3 Delays as the number of vehicle increases (*n*) on a single road segment

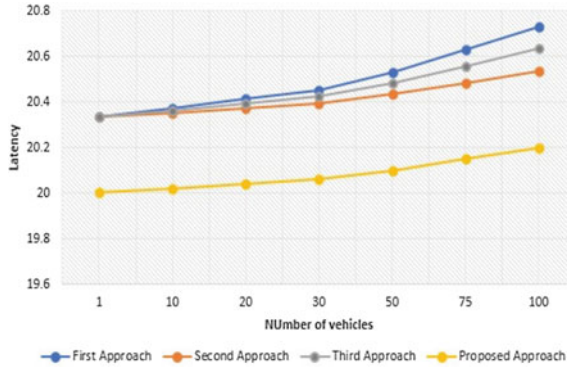| Delays	CRATER	Aid Drivers	Vehicle with Sensors	Proposed approach
Number of packets	*N*	*n*	*n*	*n*
Processing delay (total)	2 * *n*	*n*	2 * *n*	0.001
Queuing delay	*N*	*n*	*n*	*n*
Transmission delay	20	20	20	20
Propagation delay	0.3336	0.3336	0.3336	0.00003336
Traffic	*N*	*n*	*n*	*n*
Latency	20.3356	20.3346	20.3356	20.00103336
Throughput	0.049174846	0.049177264	0.049174846	0.049997417

Fig. 13 Latency as the number of vehicle increases on a single road

Fig. 14 Throughput as the number of vehicle increases on a single road
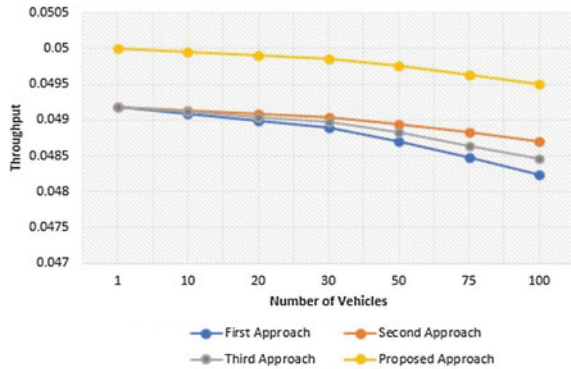
approach has low latency and high throughput. Reduced latency is a very crucial thing in VANET, where speed is critical. A split-second delay in response can cause a loss of a person's life. The proposed approach has significantly decreased this delay. Existing approaches use the cloud-based server to store the data; hence, they rely on internet connectivity and would be unable to operate in areas with poor internet access. In contrast, the proposed approach stores the data on the local server of the RSU, and hence, it does not require internet access. So proposed approach is faster and more reliable compared to the other existing approach. RSUs are already installed on the road segments in different forms such as a tower, a signpost, or anything else. So the real hardware cost is simply delivering memory to these RSUs. Sensors such as LiDAR could be an excellent alternative for both detection and spatial location.

Table 4 Delays as the number of road increases (vehicles per road remain constant)

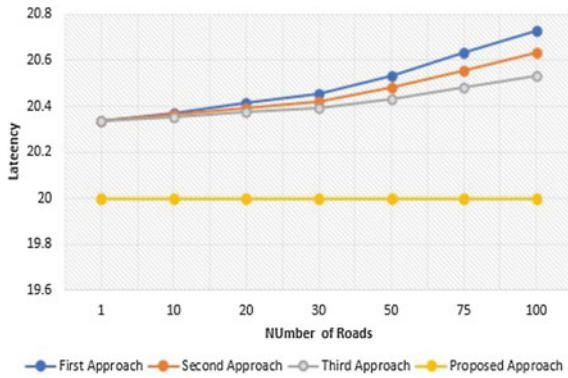| Delays	Crater	Aid drivers	Vehicle with sensors	Proposed approach
Number of packets	n	N	n	n
Processing delay (total)	$2 * n$	N	$2 * n$	0
Queuing delay	n	N	n	0
Transmission delay	20	20	20	20
Propagation delay	0.3336	0.3336	0.3336	0.00003336
Traffic	n	N	n	0
Latency	20.3356	20.3346	20.3356	20.00103336
Throughput	0.049174846	0.049177264	0.049174846	0.049997417

Fig. 15 Latency as number of roads increases and vehicles per road remain constant
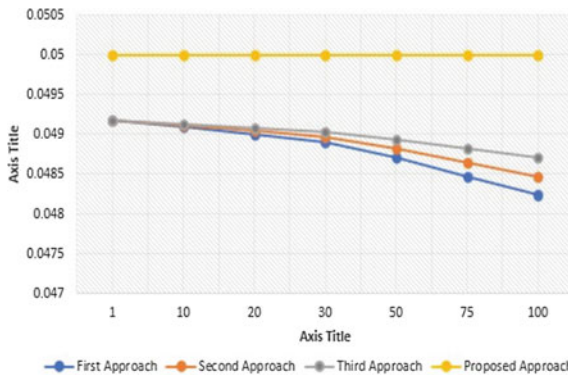

Fig. 16 Throughput as number of roads increases and vehicles per road remain constant

6 Conclusion and Future Scope

Potholes are a major issue, especially in countries like India, where they are common due to poor road construction and maintenance. In this paper, we have proposed a pothole detection and warning system for autonomous vehicles. This system notifies the vehicle in advance about the upcoming pothole and hence can avoid accidents. The results of the simulation have proved that the proposed approach is better than the existing counterparts as here the data is stored locally on the RSU. Local storage on the RSU decreases response time and delay and increases throughput. The proposed approach is independent of internet connectivity; thus, it reduces network congestion and hence improves reliability. The proposed approach accurately detects and notifies of the upcoming pothole reducing the number of accidents and saving thousands of priceless lives.

As a future scope, we will employ big data approaches in this scenario to tackle the huge data that accumulates on the RSUs, so that the information becomes more accurate, resulting in improved results and efficiency. Also, currently, we have used simulation to obtain the desired results; in the future, we will try to implement the proposed approach to prove that the results obtained through simulation are significant.

References

1. Ministry of road transport and highway transport research wing. Road Transport Yearbook 2017–18 2018–19 (2021). https://morth.nic.in/sites/default/files/RTYB-2017-18-2018-19.pdf [Online; Accessed 1 Jun 2022]
2. Sam D, Evangelin E, Raj VC (2015) A novel idea to improve pedestrian safety in black spots using a hybrid VANET of vehicular and body sensors. In: International conference on innovation information in computing technologies, pp 1–6
3. Transport department of government of Jharkhand. Causes of road accidents. https://jhtransport.gov.in/causes-of-road-accidents.html [Online; Accessed 25-May 2022]
4. HT digital streams limited. Over 3.54lakh road accidents in 2020, more than 60% due to over speeding (2016). https://www.livemint.com [Online; Accessed 1 June 2022]
5. HT Digital Streams Ltd. India loses $58 billion annually due to road accidents (2016). https://www.livemint.com/Politics/F9ljlqoWYdxxgJZ4razuiI/India-loses-58-billion-annually-due-to-road-accidents-UN-s.html [Online; Accessed 1 June 2022]
6. Surya Narayana G, Kolli K, Ansari MD, Gunjan VK (2021) A traditional analysis for efficient data mining with integrated association mining into regression techniques. In: ICCCE 2020, Springer, Singapore, pp 1393–1404
7. Shad Withers. Vehicle Damage Due To Poor Road Conditions (2022). https://www.nolo.com/legal-encyclopedia/vehicle-damage-due-to-poor-road-conditions-who-is-liable.html [Online; Accessed 2 June 2022]
8. The national center for families learning. What is a pothole?. https://www.wonderopolis.org/wonder/what-is-a-pothole [Online; Accessed 2 June 2022]
9. Jobayer A, Masud AI, Sharin ST, Shawon KFT, Zaman Z (2021) Pothole detection using machine learning algorithms. In: 2021 15th International conference on signal processing and communication systems (ICSPCS) pp 1–5

10. Kashyap A, Gunjan VK, Kumar A, Shaik F, Rao AA (2016) Computational and clinical approach in lung cancer detection and analysis. Procedia Comput Sci 89:528–533
11. Aditya, from Symbiosis Law School. Pothole Deaths in India (2021). https://blog.ipleaders.in/pothole-deaths-india/ [Online; Accessed 1 June 2022]
12. Hamza Tanveer, LaibaIkram. Pothole detection using an Android Application (2021). https://sites.google.com/view/mpvir/projects/pothole-detection-using-an-android-application [Online; Accessed 1 June 2022]
13. Vishal Khanna. Killer Potholes On Indian Roads (2019). https://gomechanic.in/blog/potholes-on-indian-roads/ [Online; Accessed 1 June 2022]
14. Kumar S, Ansari MD, Gunjan VK, Solanki VK (2020) On classification of BMD images using machine learning (ANN) algorithm. In: ICDSMLA 2019, Springer, Singapore, pp 1590–1599
15. Singal G, Goswami A, Gupta S, Choudhary T (2018) Pitfree: potholes detection on Indian roads using mobile sensors. In: 2018 IEEE 8th international advance computing conference (IACC), pp 185–190
16. Silvister S, Komandur D, Kokate S, Khochare A, More U, Musale V, Joshi A (2019) Deep learning approach to detect potholes in real-time using smartphone. In: 2019 IEEE pune section international conference (PuneCon), pp 1–4
17. Li Y, Papachristou C, Weyer D (2018) Road pothole detection system based on stereo vision. In: NAECON 2018—IEEE national aerospace and electronics conference, pp 292–297
18. Prasad PS, Sunitha Devi B, Janga Reddy M, Gunjan VK (2018) A survey of fingerprint recognition systems and their applications. In: International conference on communications and cyber physical engineering, Springer, Singapore, pp. 513–520
19. Syfullah M, Lim JMY (2017) Data broadcasting on cloud-VANET for IEEE 802.11p and LTE hybrid VANET architectures. In: 2017 3rd international conference on computational intelligence communication technology (CICT), pp 1–6
20. Kanchan S, Chaudhari NS (2018) SRCPR: SignReCrypting proxy resignature in secure VANET groups. IEEE Access 6:59282–59295
21. Hu S, Jia Y, She C (2017) Performance analysis of VANET routing protocols and implementation of a VANET terminal. In: 2017 International conference on computer technology, electronics and communication (ICCTEC), pp 1248–1252
22. Ahmed SM, Kovela B, Gunjan VK (2020) IoT based automatic plant watering system through soil moisture sensing—a technique to support farmers' cultivation in Rural India. In: Advances in Cybernetics, Cognition, and machine learning for communication technologies, Springer, Singapore, pp 259–268
23. Rashid E, Ansari MD, Gunjan VK, Ahmed M (2020) Improvement in extended object tracking with the vision-based algorithm. In: Modern approaches in machine learning and cognitive science: a walkthrough, Springer, Cham, pp 237–245
24. Kaur R, Singh TP, Khajuria V (2018) Security issues in vehicular adhoc network(VANET). In: 2018 2nd international conference on trends in electronics and informatics (ICOEI), pp 884–889
25. Saini M, Alelaiwi A, Saddik AE (2015) How close are we to realizing a pragmatic VANET solution? a meta-survey. ACM Comput Surv 48:1–40
26. Kaur H, Meenakshi (2017) Analysis of VANET geographic routing protocols on real city map. In: 2017 2nd IEEE international conference on recent trends in electronics, information communication technology (RTEICT) pp 895–899
27. Vijayakumar V, Inbavalli P, Joseph KS, Amudhavel J, Rajaguru D, Kumar SS, Vengattaraman T, Premkumar K (2015) Quantitative analysis on various safety centric based approaches in VANET. In: 2015 Global conference on communication technologies (GCCT), pp 834–837
28. Singh N, Ahuja NJ (2019) Implementation and evaluation of intelligence incorporated tutoring system. Int J Innovative Technol Exploring Eng 8(10C):4548–4558
29. Farrokhi G, Zokaei S (2014) Improving safety message dissemination in IEEE 802.11e based VANETs using direction oriented controlled repetition technique. In: 2014 IEEE 21st symposium on communications and vehicular technology in the Benelux (SCVT), pp 100–104

30. Legal Information Institute (LII). Definitions, OBUs. https://www.law.cornell.edu/cfr/text/47/95.3103 [Online; Accessed 1 June 2022]
31. Li H, Pei L, Liao D, Chen S, Zhang M, Xu D (2020) Fadb: A finegrained access control scheme for VANET data based on blockchain. IEEE Access 8:85190–85203
32. Farouk F, Alkady Y, Rizk R (2020) Efficient privacy-preserving scheme for location based services in VANET system. IEEE Access 8:60101–60116
33. Sahu H, Singh N (2018) Software-defined storage. In: Innovations in software-defined networking and network functions, IGI Global, pp 268–290
34. Li H, Song D, Liu Y, Li B (2019) Automatic pavement crack detection by multi-scale image fusion. IEEE Trans Intell Transp Syst 20(6):2025–2036
35. Shi Y, Cui L, Qi Z, Meng F, Chen Z (2016) Automatic road crack detection using random structured forests. IEEE Trans Intell Transp Syst 17(12):3434–3445
36. Xu G, Ma J, Liu F, Niu X (2008) Automatic recognition of pavement surface crack based on BP neural network. In: 2008 International conference on computer and electrical engineering, pp 19–22
37. Mishra B, Singh N, Singh R (2014) Master-slave group based model for co-ordinator selection, an improvement of bully algorithm. In: 2014 International conference on parallel, distributed and grid computing. IEEE, pp 457–460
38. Zhang L, Yang F, Zhang YD, Zhu YJ (2016) Road crack detection using deep convolutional neural network. In: 2016 IEEE international conference on image processing (ICIP), pp 3708–3712
39. Zhu S, Xia X, Zhang Q, Belloulata K (2007) An image segmentation algorithm in image processing based on threshold segmentation. In: 2007 3rd International IEEE conference on signal-image technologies and internet-based system, pp 673–678
40. Lakshmanna K, Shaik F, Gunjan VK, Singh N, Kumar G, Shafi RM (2022) Perimeter Degree technique for the reduction of routing congestion during placement in physical design of VLSI circuits. Complexity
41. Su H, Maji S, Kalogerakis E, Learned-Miller E (2015) Multi-view convolutional neural networks for 3d shape recognition. In: 2015 IEEE Int Conf Comput Vis (ICCV), pp 945–953
42. Fan R, Ozgunalp U, Hosking B, Liu M, Pitas I (2020) Pothole detection based on disparity transformation and road surface modeling. IEEE Trans Image Process 29:897–908
43. Cao W, Liu Q, He Z (2020) Review of pavement defect detection methods. IEEE Access 8:4531–14544
44. Hegde S, Mekali HV, Varaprasad G (2014) Pothole detection and inter vehicular communication. In: 2014 IEEE international conference on vehicular electronics and safety, pp 84–87
45. Kalim F, Jeong JP, Ilyas MU (2016) Crater: a crowd sensing application to estimate road conditions. IEEE Access 4:8317–8326
46. Singh N, Ahuja NJ (2019) Bug model based intelligent recommender system with exclusive curriculum sequencing for learner-centric tutoring. Int J Web-Based Learn Teach Technol (IJWLTT) 14(4):1–25
47. Madli R, Hebbar S, Pattar P, Golla V (2015) Automatic detection and notification of potholes and humps on roads to aid drivers. IEEE Sens J 15(8):4313–4318
48. Edwan E, Sarsour N, Alatrash M (2019) Mobile application for bumps detection and warning utilizing smartphone sensors. In: 2019 International conference on promising electronic technologies (ICPET), pp 50–54
49. Singh N, Kumar A, Ahuja NJ (2019) Implementation and evaluation of personalized intelligent tutoring system. Int J Innov Technol Explor Eng (IJITEE) 8:46–55
50. Basavaraju A, Du J, Zhou F, Ji J (2020) A machine learning approach to road surface anomaly assessment using smartphone sensors. IEEE Sens J 20(5):2635–2647
51. Lekshmipathy J, Velayudhan S, Mathew S (2020) Effect of combining algorithms in smartphone based pothole detection. Int J Pavement Res Technol 14:07
52. Carlos MR, González LC, Wahlström J, Cornejo R, Martínez F (2021) Becoming smarter at characterizing potholes and speed bumps from smartphone data—introducing a second-generation inference problem. IEEE Trans Mobile Comput 20(2):366–376

53. Johari S, Krishna MB (2021) TDMA based contention-free MAC protocols for vehicular ad hoc networks: a survey, vehicular communications 28(100308), ISSN 2214-2096. https://doi.org/10.1016/j.vehcom.2020.100308. https://www.sciencedirect.com/science/article/pii/S2214209620300796
54. Johari S, Krishna MB (2022) Time-slot reservation and channel switching using markovian Model for Multichannel TDMA MAC in VANETs. In IEEE Access 10:81250–81268. https://doi.org/10.1109/ACCESS.2022.3196031
55. Johari S, Krishna MB (2022) Prioritization for time slot allocation and message transmission in TDMA MAC for VANETs. In 2022 IEEE 11th International conference on communication systems and network technologies (CSNT), pp. 515–520. https://doi.org/10.1109/CSNT54456.2022.9787621