Lecture Notes in Electrical Engineering 789

N. R. Shetty · L. M. Patnaik ·
H. C. Nagaraj · Prasad N. Hamsavath ·
N. Nalini *Editors*

# Emerging Research in Computing, Information, Communication and Applications

## ERCICA 2020, Volume 1

Springer

# Lecture Notes in Electrical Engineering

## Volume 789

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

**China**

Jasmine Dou, Editor (jasmine.dou@springer.com)

**India, Japan, Rest of Asia**

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

**Southeast Asia, Australia, New Zealand**

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

**USA, Canada:**

Michael Luby, Senior Editor (michael.luby@springer.com)

**All other Countries:**

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at http://www.springer.com/series/7818

N. R. Shetty · L. M. Patnaik · H. C. Nagaraj ·
Prasad N. Hamsavath · N. Nalini
Editors

# Emerging Research in Computing, Information, Communication and Applications

ERCICA 2020, Volume 1

Springer

*Editors*
N. R. Shetty
Nitte Meenakshi Institute of Technology
Bengaluru, Karnataka, India

L. M. Patnaik
National Institute of Advanced Studies
Bengaluru, Karnataka, India

H. C. Nagaraj
Nitte Meenakshi Institute of Technology
Bengaluru, Karnataka, India

Prasad N. Hamsavath
Nitte Meenakshi Institute of Technology
Bengaluru, Karnataka, India

N. Nalini
Nitte Meenakshi Institute of Technology
Bengaluru, Karnataka, India

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-16-1337-1 ISBN 978-981-16-1338-8 (eBook)
https://doi.org/10.1007/978-981-16-1338-8



This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

# Contents

# About the Editors

**Prof. N. R. Shetty** is Chancellor of Central University of Karnataka, Kalaburagi, and Chairman of the Review Commission for the State Private University Karnataka. He is currently serving as an advisor to the Nitte Meenakshi Institute of Technology (NMIT), Bengaluru. He is also Founder Vice-President of the International Federation of Engineering Education Societies (IFEES), Washington DC, USA. He served as Vice Chancellor of Bangalore University for two terms and President of the ISTE, New Delhi three terms. He was also Member of the Executive Committee of the AICTE and Chairman of its South West Region Committee.

**Prof. L. M. Patnaik** obtained his Ph.D. in 1978 in the area of Real-Time Systems and D.Sc. in 1989 in the areas of Computer Systems and Architectures, both from the Indian Institute of Science (IISc), Bengaluru. During March 2008 to August 2011, he was Vice Chancellor, Defence Institute of Advanced Technology, Deemed University, Pune. Currently, he is Honorary Professor with the Department of Electronic Systems Engineering, Indian Institute of Science, Bengaluru, and INSA Senior Scientist and Adjunct Professor with the National Institute of Advanced Studies, Bengaluru. During the last 50 years of his long service, his teaching, research, and development interests have been in the areas of parallel and distributed computing, computer architecture, CAD of VLSI systems, high-performance computing, mobile computing, theoretical computer science, real-time systems, soft computing and computational neuroscience including machine cognition. In these areas, he has 1286 publications in refereed international journals and refereed international conference proceedings including 30 technical reports, 43 books and 26 chapters in books.

**Dr. H. C. Nagaraj** was born on 4th April, 1959 in Bengaluru. He obtained his B.E. Degree in Electronics and Communication Engineering securing first class with distinction from the University of Mysore in 1981. He obtained his M.E. in Communication Systems from P. S. G. College of Technology, Coimbatore, Bharathiyar University in 1984 and secured first rank. He got his doctoral degree in the area of

Biomedical Signal Processing and Instrumentation from Indian Institute of Technology Madras, Chennai in 2000. He assumed charge as Principal, Nitte Meenakshi Institute of Technology, Bengaluru in September 2003 and till date he is serving as Principal. He has about 38 years of experience in teaching, research and administration. During his term as Principal in NMIT, Bengaluru, the institution has been conferred with Autonomous Status by the Visvesvaraya Technological University, Government of Karnataka and UGC, New Delhi; accredited by the National Assessment and Accreditation Council, National Board of Accreditation, AICTE, New Delhi for UG programmes, and a pico satellite weighing less than 1 kg was designed and fabricated and handed over to ISRO and was launched on 12th July 2010 from Sriharikota, Andhra Pradesh, India. Dr. H. C. Nagaraj has published a number of papers in National and International Journals, Conferences and delivered invited talks in the field of Biomedical Signal Processing, Image Processing, Mobile Communication etc. He got the Best Paper Award in the National Conference of Biomechanics held during February 22–24, 1996 at IIT Madras. He was awarded the Best Paper Award at State Level Seminar on "Introduction of Flexible System in Technical Education" held by the Government of Karnataka during February 1999 in Bengaluru. He has guided several students for their Ph.D. programmes and written a book titled *VLSI Circuits* published in 2006 for B.E. Students. He is a member of expert teams visiting various colleges for accreditation by NAAC and affiliation of UG, PG and Ph.D. programmes under Electronics and Communication, Biomedical Engineering Streams, on behalf Institution of Engineers, Visvesvaraya Technological University, Belgaum and Kuvempu University, Shimoga. He is a Fellow of the Institution of Electronics and Telecommunication Engineers (IETE), Life member of ISTE, Life Member of Biomedical Engineering Society of India and Member of Global Engineering Deans Council India Chapter (GEDCIC). Dr. Nagaraj has visited countries like USA, Canada, UK, Singapore, Malaysia, South Korea, Hungary etc. in connection with collaboration for academic activities. Presently, he is the Member of the Academic Senate at Visvesvaraya Technological University (VTU), Belagavi. He is also the Member of Court at Pondicherry University and Member of Expert Team for the evaluation of R&D projects funded by VTU.

**Dr. Prasad N. Hamsavath** is currently working as Professor and Head of the Department of Master of Computer Applications at Nitte Meenakshi Institute of Technology, Bengaluru. He completed his Ph.D. at Jawaharlal Nehru University, New Delhi, India. Dr. Prasad has more than 12 years of experience in different roles in both public and private sector enterprises, including the Ministry of Human Resource and Development, New Delhi, Government of India. He has received the prestigious "Dr. Abdul Kalam Life Time Achievement Award" and also received a "Young Faculty" award at the 2nd Academic Brilliance Awards.

**Dr. N. Nalini** is currently working as Professor at the Department of Computer Science and Engineering at Nitte Meenakshi Institute of Technology, Bengaluru. She received her M.S. from BITS, Pilani, in 1999, and her Ph.D. from Visvesvaraya

Technological University in 2007. She has more than 17 years of teaching and 10 years of research experience. She has several international publications to her credit, and has received the “Bharath Jyothi Award” by Dr. Bhishma Narain Singh, Former Governor of Tamil Nadu and Assam, given out by the India International Friendship Society. She is Lifetime Member of the ISTE, CSI, ACEEE and IIFS.

# Design of a Secure Blockchain Based Privacy Preserving Electronic Voting System

**R. Shashidhara, M. Indushree, and N. S. Sneha**

## 1 Introduction

Voting is the foundation of any successful democracy and must therefore be accessible and secure for all eligible citizens in the country. Several Electoral systems take on to permit citizens to cast their precious vote, which includes electronic methods, ballot based voting and Electronic Voting Machine (EVM). However, we argue that existing techniques for voting, based on electronic voting machines, provides mistrust kind of transparency to voters. The issue commonly known as voter confidence. The Voting Systems have to heighten privacy and secrecy to provide electoral services available to the voters but secured against security vulnerabilities like keeping the voter ballot from being modified with the impact of changing casted votes by the voter. Several voting machines depends on Tor to provide anonymity of voters. Nevertheless, this mechanism doesn't achieve voter privacy and integrity services. Because, most of the intelligence authorities in the world is controlled by various parts of the Internet, which leads attackers to eavesdrop votes. As a result, as an alternative of move back to an inefficient and traditional mechanisms, the modernization of state structure by the make use of emerging technology like Blockchain [9].

Blockchain is a digital public ledger that records online transactions. Blockchain ensure security services like confidentiality, integrity, privacy by encrypting and

R. Shashidhara (✉) · M. Indushree
School of Engineering and Applied Sciences, Bennett University, Greater Noida, Uttar Pradesh, India

M. Indushree
e-mail: E20SOE815@bennett.edu.in

N. S. Sneha
Department of Computer Science and Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Udupi 574115, Karnataka, India
e-mail: snehans.cs@sode-edu.in

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_1

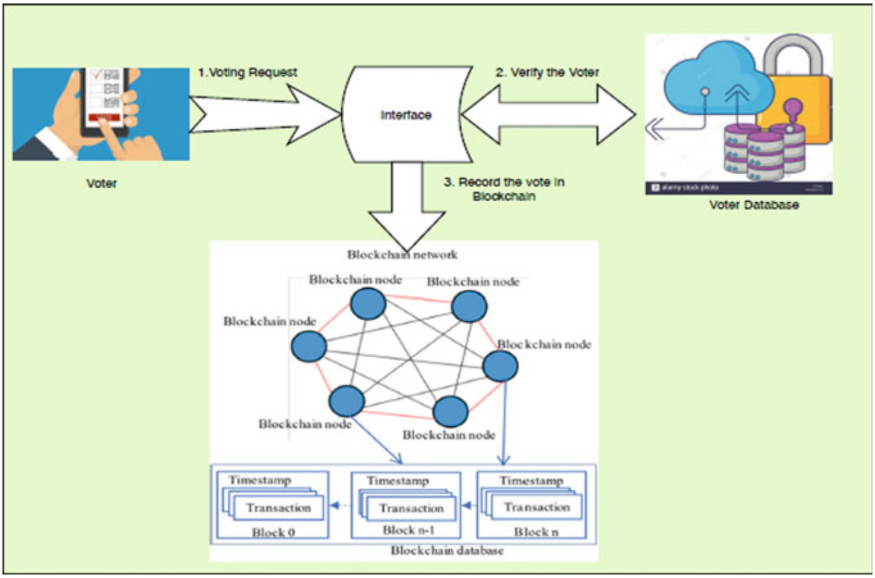



**Fig. 1** The scenario of blockchain based E-Voting system

validating the transactions. In the Blockchain, when new block is added it will be connected to the last block using a cryptography hash produced from the information of the previous block, which assure that the chain in the blockchain is never broken and blocks are permanently stored. Further, it is highly impossible to modify previous transactions. Because, all the adjacent blocks must be modified first. This fundamental aspect of Blockchain is what makes the technology tamper-proof and secure. The scenario of Blockchain based E-Voting system is depicted in Fig. 1.

## *1.1 Motivations*

Controversial E-Voting could have been avoided if the election and counting process is transparent, verifiable and secure. The existing voting system does offer voter privacy and even the vote counting by the officials is also not transparent. The voters are supposed to trust the result which is provided by the government body or Election Commission. There are also other electoral flaws like ballot stuffing, voter fraud and booth capturing. These issues makes difficult for election commission to differentiate between real votes and votes added without proper authentication and authorization. Some of the problems in current Electoral process are listed below:

1. Designing a secure electronic voting system that offers the transparency, privacy and fairness has been a challenging.

2. Current E-Voting protocols require a centralized authority to monitor and control the whole procedure from ballot to results.
3. The centralized systems are vulnerable to security attacks like Distributed Denial-Of-Service (DDOS).
4. Intelligence agencies have access to network and sufficient computing resources to analyze voting informative for the potential modification.

### *1.2 Research Contributions*

1. Designing a secure and decentralized Blockchain based E-Voting system using smart contracts (Chain code).
2. A user credential model will be proposed to ensure authentication, authorization and non-repudiation services.
3. The voter can cast a vote using private key, after that transaction will be recorded in the decentralized Blockchain network.
4. With help of voter Ethereum address, he/she can verify the casted vote in the later stages.
5. Further, candidate count with details, Vote Count and winning proposal is implemented using smart contracts and deployed in Blockchain network.

### *1.3 Structure of the Article*

The rest of the article is structured as follows: Sect. 2, covers background, it includes challenges of the E-Voting and Blockchain. Section 3, defines the security requirements of E-Voting. Section 4, describes the proposed Blockchain based E-Voting protocol. Section 5, provides the implementation details with the experimental results and provides a performance analysis of the proposed work. Section 6, concludes the article.

## 2 Background

Mistrust in the E-Voting process is a common circumstance even in the developed countries. To ensure the transparency and security, that can be implemented in Blockchain based E-voting system. This could helps in solving most of the issues being faced in the voting process.

The concept of E-voting was initiated in 2001 at Estonia. They use digital smart cards for identification and authentication. For voters to attain the voting process by displaying contestants and start casting the votes through portal in the web as well as similar desktop application. In this regard, anyone having the smart device with

Internet connection and ID card by the government, can easily vote from anywhere [5].

The E-voting is based on centralized solution have a single point of failure, which leads to security attacks and vulnerabilities. In this context, Denial of Service, Man-In-The-Middle and insider attacks could crash the centralized databases and servers. The admins of the systems could act malicious and manipulate the information [2].

Brennan Center for Justice in 2015 identified the security vulnerabilities in America voting machines and published in the news. The study identified that, 43 out of 50 US states used Electronic Voting Machines (EVM) that are old voting equipment's, exposed to crashes and failures. Further, the EVMs also easy to crack and modify with [8].

Zhao et al. presented an E-voting scheme, which proposes the reward and penalty based protocol for safe or unsafe conduct of voters. Notably, this is a first Blockchain based voting system [7]. Additionally, in 2016, Lee et al. introduced the voting scheme, which includes a trust third party using Blockchain to ensure choice of the voters. Although, the authors Bistarelli et al. presented an E-voting system, which partitions the voting process into two different parts called authentication and Distribution Server using token to safeguard privacy of the voters. Nevertheless, still there are some problems in this E-Voting scheme [1]. This system have been used in the countries like Ireland, Norway and Estonias [3]. Recently, there have been scenarios where it was faced several issues like transparency, fairness and not completely hygienic, which can be identified in countries like Brazil, Nigeria, India, Bangladesh and Pakistan [4, 6]. Notably, The issues causing the mistrust in the voting process are listed in Table 1.

**Table 1** Nature of problems causing mistrust in the E-Voting process

| Issues | Description |
|---|---|
| Casting duplicate votes | If there is no proper authentication and authorization, it is possible to cast again for the ones who have not voted |
| Pre-poll rigging | In few places the polling stations are made too far and voters have no interest or refuse to vote |
| Use of power to influence | The use of power to influence the voters or polling staff either by threats or by incentives based |
| Lack of interest by public | Voters are not fully trusted and convinced with current voting system. These issues can be dealt with trustworthy E-Voting platforms like Blockchain |
| Unsupervised vote counting | For the parties who do not have a strong representation in a region, it is likely their votes can be miscounted |
| Lack of audit and appeals | The process of hearing and deciding the appeals on some issues is slow that can be finalized before the next elections |

## 3 Security Requirements for E-Voting

The proposed Blockchain based E-voting protocol should satisfy the following security requirements:

1. Eligibility: The authorized voters should be allowed to participate in voting process and cast their vote only once in the election. Further, the system must validate the voter identities.
2. Voter Privacy: E-Voting protocol should not reveal the identities of the voter and not establish any links between identity information and ballots. Participants should remain anonymous and voting information is untraceable during and after the election.
3. Fairness: No election results should be leaked before completion of the election process. This ensures that the voters might not be affect by others in the voting process.
4. Verifiability: This security service assure that all entities in E-Voting should have the facility to verify whether the vote casted have been counted or not. Here, an individual verifiability gives the voter to verify that one's vote has been counted.
5. Forgiveness: The ability of the voter to modify ones vote after it has been cast.

## 4 Proposed System

The motive beyond the proposed mechanism is to have the Blockchain based system that satisfies the mentioned security requirements and goals. The proposed system has been designed to achieve the high degree of decentralization to create the system which the voter reign as the network of nodes.

The first transaction added to the blockchain will represent the genesis block. When a voter cast his/her vote, the transaction is updated in the Blockchain network. The proposed e-Voting protocol permit for the protest vote, where an user might be return the blank vote to the refusal of the election system or like NOTA to dissatisfaction with all candidates. The Blockchain is decentralized Peer-to-Peer network and cannot be immutable. Even there is no central point of failure. In order to ensure the security and trust, the current block will uses the previous block hash like the previous voters data. If any of the blocks are corrupted, modified then it will be effortless to trace out. Because, all blocks in the blockchain are linked to each other with previous hash and serves as chain. During voting in the Blockchain, the vote gets transmitted to the nodes on Blockchain network. After that the node adds vote to the decentralized network.

The proposed protocol consists of the following phases:

1. **Setup**: This is an initialization phase to obtain the private key and public key pair using asymmetric cryptosystem.

2. **Voter Authentication**: The user should logs to the system using the credentials. The protocol will authenticate the voter based on his/her identity information issued by the Election Commission. The E-voting system should verify and validate all information entered by the voter. If the verification is successful, the voter will be authenticated and authorized to cast the vote.
   The prototype for voter authentication is described below:
   - **(ID, PW)**: Enter the login details and link the node identity to the e-governance.
   - **(Credentials, node id, user-Info)**: The system of E-governance authenticates voter credentials.
3. **Casting a vote:** Voters should choose the candidates from list of contestants to cast their vote. The voter can cast the vote through a friendly user interface. The prototype for this phase is shown below:
   - **V = vote (ID of the voter, candidate selected).**
   - **Add (V, Chain)**, the Vote V is added to the Blockchain network.
   - Next, the updated Blockchain data is reacted in all the nodes.
   - **Vote (ID, user List, true)**: Finally, Voter field will be switched to vote.
4. **Formation of the Block:** Upon casting the vote by the voter will be recorded as a unconfirmed transaction in the Blockchain. The nodes in the Blockchain network will validate the casted vote based on consensus protocols.
5. **Sealing of Blocks:** The transactions are stored in the Blockchain, by the end of polling time all blocks in the network needs to be sealed by cryptographic hash (SHA-256) using nonce and merkle root. Once the electoral process is complete and the results have been published, then there is no significance for the Blockchain mining.
6. **Counting of votes:** We have implemented a mechanism to count the casted votes in a fair manner. Further, the proposed system supports the voter to check the casted vote is successfully counted during counting process or not. With help of the Ethereum address, a user can verify the status of the casted vote.
   The prototype for counting process is given below:
   - **Candidates = get Candidates (candidate List).**

   Receive the candidate details from E-Governance.
   - **Results = count (chain, candidates)**

   Here, vote counting process will be completed and winner will be identified based on the maximum number of votes.

## 5 Implementation and Experimental Evaluation

The proposed E-voting protocol is implemented using Ethereum platform called public Blockchain network. An Ethereum network provides a broader range of applications, with the power of smart contracts. Ethereum Blockchain consists of Ethereum nodes. The node is any device that is running the Ethereum protocol (blockchain). When we connect to the Ethereum protocol we are on the Ethereum Blockchain network. By running an Ethereum node we can connect to other nodes in the network, have direct access to the Blockchain, and even do things like mine blocks, send transactions, and deploy smart contracts. Many applications, that may normally require a web server, can be run through these smart contracts using Blockchain network. Hence, it is impossible to manipulate the transactions or smart contracts deployed in the Blockchain network.

After performing the transactions on Ethereum Blockchain network, the transaction fee is calculated in Gas, and paid for in Ether. The gas is the fuel of the Ethereum network, which is mainly used to conduct transactions, execute smart contracts and Launch Decentralized Applications (Dapps). The frequently used parameters in the Ethereum network are Gas, Gas price and Gas limit.

1. Ether (ETH) is the Ethereum network's native cryptocurrency, the second largest by market cap on the crypto market.
2. **Gas**: is the unit of calculation that indicates the fee for a particular action or transaction.
3. **Gas Limit**: is the maximum amount of Gas that a user is willing to pay for performing this action or confirming a transaction (a minimum of 21,000).
4. **Gas Price**: is the amount of Gwei that the user is willing to spend on each unit of Gas.

Additionally, we have set up a MetaMask wallet in order to perform the transacttions on Ethereum Blockchain network. MetaMask is just an Ethereum Browser and Ether wallet. It interacts with Ethereum Dapps and Smart Contracts without running a full Ethereum node. Furthermore, MetaMask supports to connect different Ethereum based Blockchain networks and possible to import the accounts from other accounts through private keys.

We have defined the E-voting protocol through smart contracts, which consists of programming code and stored on a Blockchain network, then it execute when certain terms and conditions are met. It is called smart because of its ability to verify and execute a contract without any help from third parties. The contract exists in the decentralized Blockchain network and contains all the terms of a particular agreement. The smart contracts are meant to provide accuracy, transparency, autonomy, security and standardization.

Smart contracts defined in solidity programming language is executed by the Ethereum nodes in the blockchain network in every 10 s, and its validated by at least by two other nodes in the blockchain network. After that, functions of contracts can be triggered and executed.

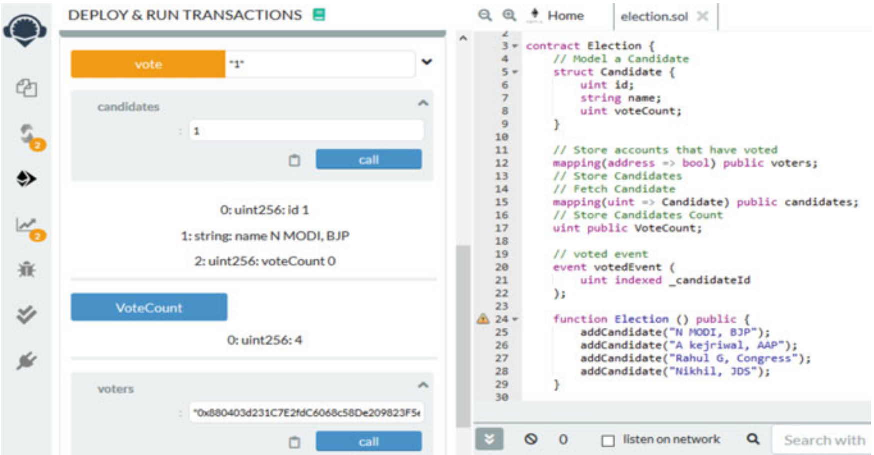



**Fig. 2** Smart contract call and execution during E-Voting

The Candidate is defined as a struct, the state variables are ID, Name and Vote-Count. We used solidity mapping for storing and fetching the voter details. ID is the wallet address associated with the voter account in the Ethereum Blockchain. The state variable VoteCount is used to count the number of votes received by the candidate.

Of the proposed protocol is evaluated by testing five ballots in the Ethereum Blockchain network. In this work, The E-Voting system scope is restricted for smaller elections and polls. The E-Voting with huge number of voters would require dynamic network structure and need to handle complex problems. The Blockchain networks scalability is still unknown. In addition, the proposed smart contracts are implemented using solidity using Ethereum platform. The wallet is supported in windows, Linux and mac machines. Furthermore, a voter who willing to cast their vote should the Ethers in his/her wallet to complete the voting transaction (Fig. 2).

## 6 Conclusion

In this article, a Blockchain based decentralized and peer-to-peer electronic voting protocol is proposed. The legitimate voters could have the power to vote through Internet by using smart devices like Mobiles, PCs, etc. The transaction will be recorded in the Blockchain network, which is verifiable, anonymous and adversaries are unable to modify the records in the network. The solidity smart contract is used to accomplish recording, managing, validating the voters during the electoral process. In order to provide the privacy and transparency of E-Voting protocol, secure cryptographic functions has been employed to ensure that the registration and voting is anonymous. The digital signatures using public key infrastructure makes the voting process more secure and reliable.

Further, the proposed protocol does not require mining like Bitcoin network since the voter's information is registered and authentic. Notably, the proposed approach addresses some of the security pitfalls that conventional E-voting protocols have. As a result of the proposed work, the concept of Blockchain technology, security algorithms and cryptographic primitives like hash functions, nonce and digital signatures, has become adaptable to elections and polls to secure the E-Voting environment.

## References


1. Akbari E, Wu Q, Zhao W, Arabnia HR, Yang MQ (2017) From blockchain to internet based voting. In: 2017 international conference on computational science and computational intelligence (CSCI). IEEE, pp 218–221
2. Ayed AB (2017) A conceptual secure blockchain-based electronic voting system. Int J Netw Sec Appl 9(3):01–09
3. Batubara FR, Ubacht J, Janssen M (2018) Challenges of blockchain technology adoption for e-government: a systematic literature review. In: Proceedings of the 19th annual international conference on digital government research: governance in the data age, pp 1–9
4. Casino F, Dasaklis TK, Patsakis C (2019) A systematic literature review of blockchain-based applications: current status, classi_cation and open issues. Telematics Inform 36:55–81
5. Hanifatunnisa R, Rahardjo B (2017) Blockchain based e-voting recording system design. In: 2017 11th international conference on telecommunication systems services and applications (TSSA). IEEE, pp 1–6
6. Hsiao JH, Tso R, Chen CM, Wu ME (2017) Decentralized e-voting systems based on the blockchain technology. In: Advances in computer science and ubiquitous computing. Springer, pp 305–309
7. Johnson D (2019) Blockchain-based voting in the US and EU constitutional orders: a digital technology to secure democratic values? Eur J Risk Regul 10(2):330–358
8. Li J, Wang X, Huang Z, Wang L, Xiang Y (2019) Multi-level multi-secret sharing scheme for decentralized e-voting in cloud computing. J Parallel Distrib Comput 130:91–97
9. Moura T, Gomes A (2017) Blockchain voting and its effects on election transparency and voter confidence. In: Proceedings of the 18th annual international conference on digital government research, pp 574–575
10. Shahzad B, Crowcroft J (2019) Trustworthy electronic voting using adjusted blockchain technology. IEEE Access 7:24477–24488
11. Yavuz E, Koc AK, Cabuk UC, Dalkilic G (2018) Towards secure e-voting using ethereum blockchain. In: 2018 6th international symposium on digital forensic and security (ISDFS). IEEE, p 17

# A Nature Inspired Algorithm for Enhancement of Fused MRI and CT Brain Images

**Leena Chandrashekar and A. Sreedevi**

## 1 Introduction

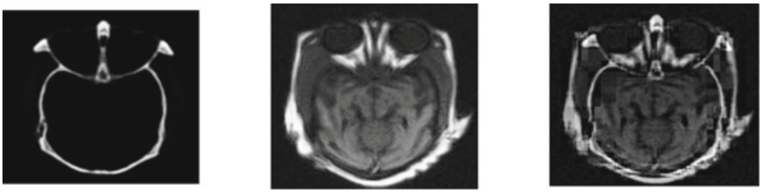
Glioblastomas are the fastest growing Grade IV malignant tumors found in the brain with a survival time of less than a year after their detection [1]. The detection of these tumors is an unceasing challenge to doctors. The failure to recognize the early symptoms, lack of awareness, inadequate imaging facilities, preliminary screening for the patients and expertise with doctors are some of the factors that delay the detection. For this reason, medical imaging is paramount in detection, identification, grading and diagnosis of the Glioblastomas. Doctors recommend various imaging techniques for detection of Glioblastomas like Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and its variants, Fluid Attenuated Inverse Recovery (FLAIR) and Positron Emission Tomography (PET) [2]. These images are acquired sequentially through different scanning machines at different times. Doctors assess the images independently and conclude the analysis based on their experience and expertise. Each of the modalities provide different information of the brain. For example, the CT image provides the structural information of the brain like bone structure, tissue symmetries, changes in tissue density and space occupying lesions [3]. It also shows changes made in the nearby skull region due to tumor extension and calcification of tumors. Conversely, CT images fail to indicate tumor borders and infiltration in the nearby regions. These can be easily visualized with Magnetic Resonance Image (MRI), which provide structural and functional information of the brain along with high contrast and resolution for soft tissues like tumors or lesions [3]. These multimodal images are the noninvasive ways to detect Glioblastomas. Based on the multimodal images, surgical resection or complete removal of tumor is made followed

L. Chandrashekar (✉) · A. Sreedevi
RVCE, Bangalore, India

A. Sreedevi
e-mail: sreedevia@rvce.edu.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_2



(a) (b) (c)

**Fig. 1** **a** CT image, **b** MR image, **c** fused image

by radio chemotherapy [4]. Thus, a single imaging technique is never sufficient to confirm the presence or extent of the tumor [5]. This helps in early detection of tumors and requires lesser digital storage [7, 8]. Some of the techniques for fusing the multimodal images are Discrete Wavelet Transform (DWT), Laplace Transform (LT), Contourlet Transform (CT) and Non-sub Sampled Contourlet transform (NSCT) [6]. The process of fusion begins with decomposing the multimodal images in to approximation coefficients (low frequency components of the image) and detailed coefficients (high frequency components) using the above techniques. These components are combined using different rules like averaging, summing, weighted summing, max–min or max-max fusion rules [7]. The fused coefficients are recomposed using inverse transforms to generate a fused image. The visual quality of the fused image is evaluated based on contrast, edge information and peak signal to noise ratio.

The fusion process introduces blocking effect, noise and artifacts that greatly reduce the visual quality of the fused image. Figure 1 shows a CT, MR and fused image obtained by DWT, indicating blocking effect at the edges and corners [6, 7]. The fused images with poor visual quality make it difficult for experts to interpret the tumor presence or its spread. Hence, there is a need for an enhancement technique to improve the visual quality and reduce blocking effect, noise and artifacts. Most of the enhancement techniques deal with improving image contrast, as it helps in differentiating the Glioblastomas and the normal cells. Nevertheless, the drawback with these techniques is that it only increases the dynamic range of the image, which is a function of pixel intensity alone. Tumors or any abnormality in the brain appears distinct, bright or light intensity in MR images. Thus, by varying the contrast, the normal and the tumorous cells can be differentiated easily. On the contrary, the high contrast images or low contrast MR image makes it challenging to differentiate the Glioblastomas. The contrast enhancement is mostly a twofold process, consisting of contrast stretch and tonal enhancement. The contrast stretch improves the brightness differences uniformly across the dynamic range of the image and tonal enhancement improve the brightness differences in different areas like dark, gray or bright regions in the image [6]. The paper deals with enhancement techniques for CT and MRI fused images that mostly focus on improving the contrast, structural information, peak signal to noise ratio with minimum loss of information.

## 2 Enhancement Techniques

Generally, the contrast enhancement is achieved by the following techniques—Non-Linear Transfer function, Histogram based and Frequency Domain [11]. Among them, Histogram Equalization (HE) is the most popular technique for contrast enhancement performed in a spatial domain. This deals with remapping the gray scale values of the original input image to a new level of gray scale values using linear or non-linear functions. This remapping aims at flattening and stretching the dynamic range of the image histogram. Although HE is a popular enhancement technique, it suffers from visual artifacts like intensity saturation and amplification due to large number of homogenous pixels. The equalization is accomplished uniformly for all the pixels of the image, leading to enhanced global contrast. However, the lowest intensity pixels become less significant, thereby reducing the local contrast [12]. In order to overcome the drawbacks of HE, newer enhancement techniques are proposed for improved visual quality with the use of median filters, adaptive gamma correction and homomorphic filtering [13, 14].

An Adaptive Histogram Equalization (AHE) technique is a block based adaptive method, that deals with the local contrast rather than global contrast. The local contrast is more significant in the detection of Glioblastomas as it can clearly distinguish the normal and tumor cells. In this technique, histogram equalization is performed on sub-images or small and equal sized blocks obtained by splitting the image. The equalization is executed on every block independently and mapped to new intensity levels based on a transformation function. The new pixel values are solely based on the neighboring pixel characteristics. Then, bilinear interpolation is used to combine the blocks after equalization [14]. The major challenge with AHE is the selection of the block size and the transformation function. AHE also suffers from blocking effect, at the time of combining the blocks. Over-amplification is also seen due to large homogenous regions of the image. Youlian Zhu et al. have proposed an adaptive histogram equalization technique for CT images. A user defined parameter $\beta$, is suggested based on the gray level of the image. The entropy is used as an objective function to select the $\beta$ adaptively [15].

A variant of AHE is the Contrast Limited Adaptive Histogram Equalization (CLAHE), proposed by *K* Zuierveld, is also a block-based contrast enhancement technique with focus on local contrast. Unlike AHE, CLAHE provides uniform equalization with clipping the excess portion of large peaks found after the histogram equalization, thereby avoiding over-amplification. The excess portion removed depends on a parameter called clip limit, which is a function of the dynamic range of the image and block size. CLAHE involves setting of three operational parameters—clip limit, block size and distribution function, which must be carefully chosen before performing the image enhancement to achieve good contrast images, free from noise and artifacts [16]. Various histogram-based enhancement techniques are compared and analyzed, CLAHE is observed to perform better for MRI brain Images [17].

The simplest technique of setting the operational parameters for CLAHE is by trial and error, however, this is time consuming, may deviate from the actual values and varies with every image. Some of the techniques used to set the parameters are based on textureness of the image, maximum curvature of entropy, Least Mean Square (LMS) algorithm, multi-objective optimization technique and fuzzy rules. In spite of various techniques for contrast enhancement, CLAHE seems to provide good local contrast, however it largely fails to enhance the pixels with low gray level intensity. Moreover, there is no standard for finding the optimal clip limit for a specific region of interest in medical images. Generally, clip limit is proportional to the multiple of mean of the histogram, where the multiplication factor is user-defined and varies for different images. Therefore, there is a need to choose clip limit adaptively for every block of image without any user intervention [23].

The clip limit is also a function of dynamic range of the grayscale image, block size and slope of transformation function. Initially, the clip limit and block size are chosen empirically and then obtain the optimal values based on statistical parameters like entropy, peak signal to noise ratio or edge information. Yet clip limit may change depending on the type of images. This makes the enhancement process very extensive and time consuming [23]. Moreover, inaccurate selection of clip limit can cause over-amplification in CLAHE. Bilateral Filter and Median filters are used to overcome this drawback [24, 25]. Optimization techniques provide a convenient way of determining the CLAHE parameters without any heuristics and compute them adaptively for every image block. Particle Swarm Optimization (PSO)—a population-based optimization technique proposed by Eberhart and Kennedy [27]. The motivation for PSO is from the biological social groupings of animals, which interact with each other to find food or save each other from predators. A swarm is defined as a group of possible solutions to the optimization problem, also called particles. Each particle of the swarm is identified with its velocity and position, which are updated through iterations. The performance of the optimization is evaluated based on a fitness function. The search for best solution terminates at the end of the iterations or when the solution generates the highest fitness value [27, 28].

Based on the recent literature, we propose an enhancement technique for fused Multimodal and Multiresolution brain images, that suffer from blocking effect, noise and artifacts during the fusion process. CLAHE can be used to improve the contrast of the fused image, but the structural information in the fused MRI and CT images is lost due to AHE. The PSO algorithm enable the selection of the operational parameters of CLAHE based on a multi-objective fitness function which is essentially dependent on entropy and edge information of the image. Hence, the enhanced image is improved in terms of contrast, structural information with minimum mean square error.

## 3 Proposed Enhancement Technique for Fused CT and MRI

The proposed enhancement technique begins with preprocessing of MRI and CT images, image registration, and image fusion followed by image enhancement. More than 200 MRI and CT images containing Grade IV tumors—Glioblastoma are taken from www.Radiopedia.org for the experiment. There are quite a few databases available publicly for MRI images, but the challenge in our research is to get multimodal images for the same patient. Since these are acquired at different times and from different machines, they must be registered and preprocessed before fusing them. The simulation is performed on MATLAB 2020. Figure 2 shows the block diagram for the enhancement process. The preprocessing stage resizes the image to size 256 $\times$ 256 and converts them to gray scale. The images contain Gaussian and Rician noises, that are eliminated; Non-Local Means Filter is used in filtering the CT and MRI images [29]. The preprocessing is followed by image registration—a mandatory step, where both the as are matched for size, orientation and scaling. Subsequently, the images are fused using Laplacian Pyramid or Non-sub sampled Contourlet transform. The former fusion technique provides good structural information and latter offers enhanced contrast in the fused images as shown in Table 1.

The fusion process introduces blocking effect and noise in the fused image, thereby reducing the image quality as discussed in Sect. 1 [30]. Since the focus is the tumor region and differentiating the healthy and tumorous cells, adaptive block-based enhancement technique like CLAHE is chosen. The operational parameters for

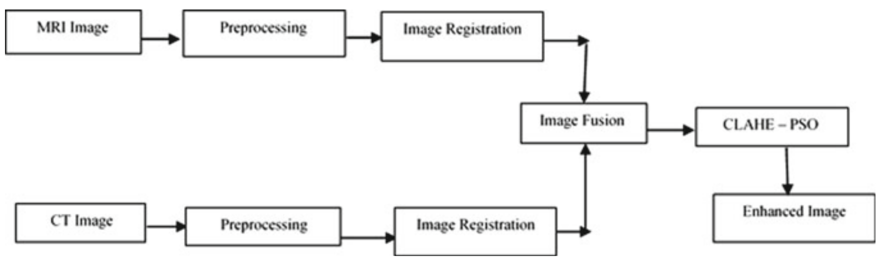



**Fig. 2** Methodology for the proposed enhancement process

**Table 1** Performance of Laplacian Pyramid and non-sub sampled contourlet transform image fusion

| Dataset D6 | STD | En | SSIM | PSNR | UIQI | MSE |
|---|---|---|---|---|---|---|
| LP Fused | 72.87 | 4.04 | 0.75 | 19.07 | 0.42 | 804.22 |
| Proposed | 68.04 | 3.93 | 0.85 | 20.4 | 0.84 | 3.44E+04 |
| NSCT fused | 76.34 | 5.81 | 0.53 | 18.88 | 0.42 | 848.33 |
| Proposed | 75.87 | 5.65 | 0.65 | 20.29 | 0.59 | 6.07E+02 |
| Dataset D13 | STD | En | SSIM | PSNR | UIQI | MSE |
| LP Fused | 72.22 | 5.39 | 0.63 | 11.92 | 0.57 | 948.58 |
| Proposed | 69.92 | 5.25 | 0.83 | 12.29 | 0.78 | 810.07 |
| NSCT fused | 77.96 | 6.24 | 0.57 | 18.61 | 0.57 | 894.85 |
| Proposed | 79.77 | 6.17 | 0.67 | 19.22 | 0.67 | 776.99 |

CLAHE are block size, clip limit and distribution function. The proposed technique is executed for different block sizes like $2 \times 2$, $3 \times 3$, $4 \times 4$, $5 \times 5$, $7 \times 7$, $8 \times 8$ and $10 \times 10$. It must be noted that, the $8 \times 8$ provided superior results in terms of contrast and structural information, hence chosen to be constant for the enhancement process. Similarly, the enhancement process was performed with the various distribution functions like Uniform, Rayleigh and Exponential distribution. The Uniform distribution allows distribution of clipped pixels unvaryingly in the image histogram and hence preferred. The clip limit is initialized randomly in the range of 0–0.01, the PSO algorithm adaptively choses the clip limit based on a fitness function, given in Eq. (1).

$$F(I_e) = \log(\log(E(I_s))) \times \frac{n\_\text{edges}(I_s)}{M \times N} \times H(I_s) \tag{1}$$

The proposed enhancement process is two-fold; firstly, the contrast of the MRI and CT fused image is enhanced by CLAHE algorithm, which increases the dynamic range of the image. Secondly, the multi-objective fitness function assists in choosing the clip limit for CLAHE that maximizes the entropy and edge information of the image. The Particle Swarm Optimization (PSO) algorithm finds the optimal clip limit. A swarm of 50 particles, each representing the clip limit is initialized in the range of 0–0.01. With each of them, the fused image is enhanced by CLAHE and a fitness function that is product of entropy, sum of edge intensities and number of edge pixels is computed. The clip limit that gives maximum fitness value is chosen. Since, multiple parameters are considered to measure the degree of enhancement; this function is called multi-objective function. Every particle $i$ in the swarm is represented by two parameters velocity and position. For any particle '$i$', the position and velocity indicate its location in the swarm and fitness value respectively. The velocity and position are computed based on some random values and is updated in every iteration using Eqs. (2) and (3).

$$v_i(t+1) = wv_i(t) + c_1 r_1 (p_i(t) - x_i(t)) + c_2 r_2 (g(t) - x_i(t)) \tag{2}$$

$$x(t+1) = x_i(t) + v_i(t+1) \tag{3}$$

where $v_i(t)$ and $x_i(t)$ represent the velocity and position for an particle $i$ and iteration $t$. Equation (2) comprises of three components—first component representing the initial velocity of the particle, the second component represents the particle's decision based on its own experience and the third component indicates the particle's decision based on swarm's experience. In every iteration, the image is enhanced using CLAHE with the selected clip limit (each particle). The fitness function is computed using Eq. (1). This process is repeated for all the particles to get the best fit (clip limit). The clip limit that maximizes the fitness function can be accessed from the swarm based on its position and velocity and is represented as '$p_{\text{best}}$' or $p_i(t)$. This denotes the best local solution for that iteration. The enhancement process is repeated for all

*Initialize the particle swarm*
*For each iteration*
*For each particle*
*Enhance the image using CLAHE*
*Compute the fitness value for the enhanced image as per Eq. (1) for every particle*
*If the fitness value is greater than the previous fitness value (pbest)*
*Set current value as the new pbest (gbest)*
*End*
*Choose the particle with the best fitness value among all the pbest (gbest)*
*For each particle*
*Calculate particle velocity as per Eq. (2)*
*Calculate the particle position as per Eq. (3)*
*End*
*Continue while maximum iterations are attained.*
*End*
*Report the gbest and pbest*

**Fig. 3** Pseudocode for proposed technique

the iterations to get '$p_{\text{best}}$' or $p_i(t)$ for each iteration. In case the '$p_{\text{best}}$' value in the current iteration is greater than the previous one, then the '$p_{\text{best}}$' is updated with a new '$p_{\text{best}}$' and '$g_{\text{best}}$', otherwise the '$p_{\text{best}}$' from previous iteration is retained as '$p_{\text{best}}$' and '$g_{\text{best}}$'. The '$g_{\text{best}}$' or $g(t)$ in Eq. (2) is the global solution for the enhancement process obtained at the end of all the iterations. When '$p_{\text{best}}$' appears equal '$g_{\text{best}}$' over a predefined number of iterations the enhancement process terminates.

A balance between '$p_{\text{best}}$' and '$g_{\text{best}}$' is achieved by inertia weight represented as $w$, $c_1$ and $c_2$—the positive acceleration constants ($c_1 = 2.5$, $c_2 = 1.5$) and $r_1$ and $r_2$ are random values in the range of [0, 1]. Figure 3 shows the pseudo code for proposed CLAHE-Particle Swarm Optimization algorithm with Uniform distribution function.

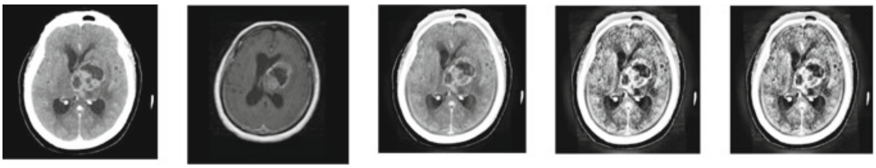
## 4 Experiments and Results

The CT and MRI images prior to enhancement are preprocessed, registered and fused using Laplacian Transform. The dataset includes many variants for CT and MRI images—(i) images with contrast agent (ii) images without contrast agent and (iii) delayed images with contrast agent. The combination of CT and MRI T1 image, CT and MRI T2 and CT and MRI FLAIR images is used to generate the fused image. The proposed enhancement technique is tested on more than 200 MRI and CT fused images, containing the different types of Glioblastomas—Multicentric Glioblastoma, Multifocal Glioblastoma, Cystic Glioblastoma and Giant Cell Glioblastoma. Table 2 lists the characteristics of the datasets considered for enhancement.

Figures 4, 5, 6 and 7 shows 4 sample CT, MRI, fused image, CLAHE enhanced image and the enhanced image from the Proposed method. The datasets chosen belong to Group I representing large tumors and Group II with small tumors. In

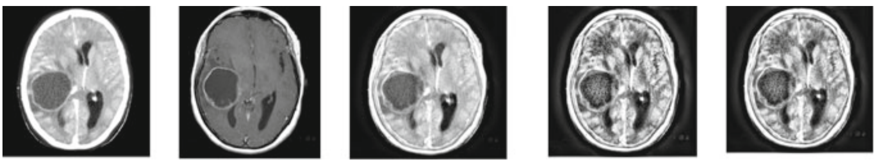
**Table 2** Characteristics for sample dataset

| Dataset | CT | MRI | Size of tumor | Group | Type of tumor |
|---|---|---|---|---|---|
| D20 | High contrast | Low contrast | Large | I | Cystic high-grade glioblastoma |
| D53 | Large size | Small size | Large | I | Glioblastoma (Grade IV) |
| D60 | Large size | Large size | Large | I | Giant cell glioblastoma |
| D67 | Large size | Small size | Large | I | Primary CNS lymphoma |
| D65 | Small size | Large size, Poor contrast | Large | I | differential diagnosis—metastasis and glioblastoma |
| D17 | Misaligned | Aligned, Poor contrast | Small | II | Glioblastoma (Grade IV) |
| D16 | Aligned | Misaligned, Poor contrast | Small | II | Multicentric glioblastoma |
| D15 | Aligned | Misaligned, Poor contrast | Small | II | Multifocal glioblastoma |
| D6 | Large | Small, Poor contrast | Small | II | Glioblastoma (Grade IV) |
| D50 | Small | Large | Small | II | Glioblastoma (Grade IV) |



(a) (b) (c) (d) (e)

**Fig. 4** Dataset D20. **a** CT image, **b** MRI image, **c** Fused image, **d** CLAHE



(a) (b) (c) (d) (e)

**Fig. 5** Dataset D53. **a** CT image, **b** MRI image, **c** Fused image, **d** CLAHE

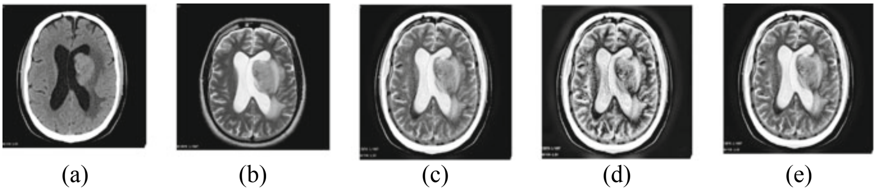


Fig. 6 Dataset D67. **a** CT image, **b** MRI image, **c** Fused image, **d** CLAHE

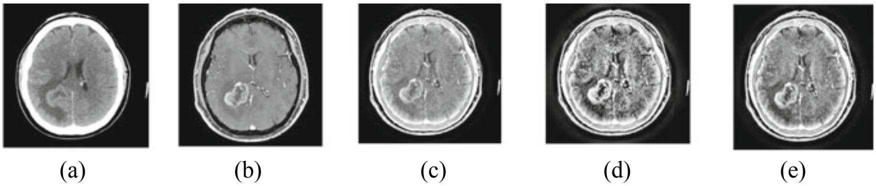


Fig. 7 Dataset D50. **a** CT image, **b** MRI image, **c** Fused image, **d** CLAHE

addition, high contrast CT and poor contrast MRI images are also tested with the proposed technique. For example, the CT image in dataset 20 (Fig. 4) has very high contrast and MRI T1 in dataset 16 and 17 has poor contrast are considered for the proposed enhancement method. Table 3 shows the expressions for various performance parameters, the detailed definitions are discussed in [30–33]. The proposed enhancement technique provides high PSNR, SSIM, UIQI and a minimum MSE. It can be observed that images enhanced by CLAHE seem enhanced, however it is only in terms of contrast or variaion of brightness but the structural information is poor. Although, the standard deviation and entropy is high for enhanced image obtained by CLAHE, the PSNR, SSIM, UIQI is lesser than those obtained with the

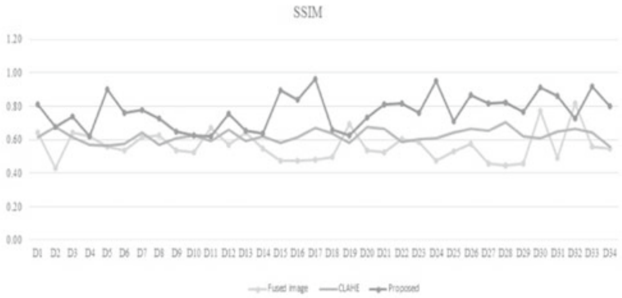
**Table 3** The performance parameters for evaluation of the enhanced image

| S. No. | Parameter | Equation |
|---|---|---|
| 1 | Standard Deviation (SD) | $\sqrt{\frac{1}{H\times W}\sum_{x=1}^{H}\sum_{y=1}^{W}(F(x,y)-\mu)^2}$ |
| 2 | Entropy (En) | $\sum_{l=0}^{L-1} p(l)\log_2 p(l)$ |
| 3 | Mean Square Error (MSE) | $\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(I_1(i,j)-E(i,j))^2$ |
| 4 | Peak Signal to Noise Ratio (PSNR) | $20\log_{10}\left(\frac{255}{\sqrt{\text{MSE}}}\right)$ |
| 5 | Structural Similarity Index Metric (SSIM) | $\frac{(2\mu_e\mu_s+C_1)(2\sigma_e+C_2)}{(\mu_e^2+\mu_s^2+C_1)(\sigma_e^2+\sigma_s^2+C_2)}$ |
| 6 | Universal Image Quality Index (UIQI) | $\frac{4\sigma_e\sigma_s(\mu_e+\mu_s)}{(\sigma_e^2+\sigma_s^2)(\mu_e^2+\mu_s^2)}$ |

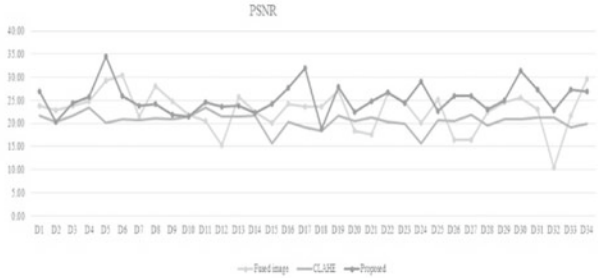
proposed enhancement technique. This is because, the CLAHE deals with enhancing the dynamic range and increasing the high intensity pixels and further decreasing the low intensity pixels of the image, however ignores the entropy or structural information of the image. Moreover, the MSE value is minimum for enhanced image in the proposed technique, indicating more similiarity to the orginal image with minimum loss of information. The clip limit in CLAHE is kept constant at 0.01, but with the proposed technique the clip limit is chosen adaptively in the range of 0–0.01 for each block of size 8 × 8, and it is found to present good results. Table 4 shows the performance paramters for the 10 image Datasets. Figure 8a–d shows the response for 34 (D1–D34) enhanced images in terms of SSIM, PSNR, UIQI and MSE.

**Table 4** Experimental results for sample dataset

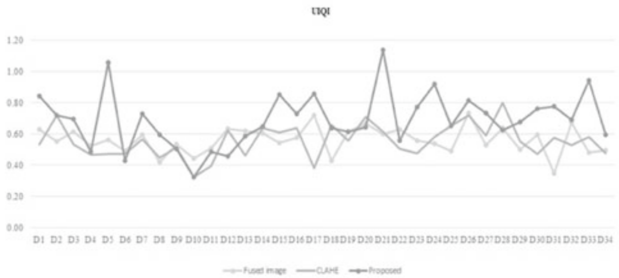
| STD | | | | En | | | SSIM | | |
|---|---|---|---|---|---|---|---|---|---|
| Dataset | Fused | CLAHE | Proposed | Fused | CLAHE | Proposed | Fused | CLAHE | Proposed |
| D20 | 91.32 | 85.24 | 85.41 | 6.51 | 6.81 | 6.78 | 0.5 | 0.70 | **0.73** |
| D53 | 84.77 | 86.03 | 85.77 | 6.24 | 6.65 | 6.49 | 0.74 | 0.65 | **0.82** |
| D67 | 70.13 | 78.01 | 71.84 | 6.57 | 7.15 | 6.75 | 0.75 | 0.65 | **0.92** |
| D65 | 83.34 | 85.03 | 83.79 | 6.32 | 6.79 | 6.32 | 0.64 | 0.69 | **0.96** |
| D60 | 72.81 | 75.03 | 73.02 | 4.92 | 5.30 | 4.96 | 0.73 | 0.59 | **0.93** |
| D17 | 79.54 | 80.34 | 80.25 | 6.23 | 6.79 | 6.78 | 0.55 | 0.62 | **0.64** |
| D16 | 78.29 | 80.19 | 79.18 | 5.60 | 6.19 | 6.11 | 0.64 | 0.59 | **0.66** |
| D15 | 84.24 | 81.39 | 81.85 | 6.28 | 7.02 | 6.69 | 0.55 | 0.60 | **0.80** |
| D26 | 74.42 | 80.23 | 77.42 | 5.30 | 5.72 | 5.54 | 0.54 | 0.57 | **0.76** |
| D50 | 73.59 | 80.09 | 75.71 | 6.22 | 6.60 | 6.34 | 0.56 | 0.64 | **0.92** |
| UIQI | | | | PSNR | | | MSE | | |
| Dataset | Fused | CLAHE | Proposed | Fused | CLAHE | Proposed | CLAHE | | Proposed |
| D20 | 0.51 | 0.73 | **0.72** | 16.22 | 19.77 | **20.86** | 685.3 | | **532.63** |
| D53 | 0.75 | 0.63 | **0.70** | 24.62 | 21.43 | **25.49** | 467.07 | | **183.57** |
| D67 | 0.72 | 0.70 | **0.95** | 16.80 | 18.17 | **27.50** | 989.00 | | **115.60** |
| D65 | 0.64 | 0.73 | **0.93** | 13.41 | 21.50 | **34.87** | 459.57 | | **21.16** |
| D60 | 0.72 | 0.33 | **0.66** | 20.13 | 23.65 | **36.39** | 280.46 | | **14.92** |
| D17 | 0.55 | 0.62 | **0.64** | 22.37 | 21.59 | **22.20** | 450.82 | | **391.61** |
| D16 | 0.62 | 0.47 | **0.59** | 25.80 | 21.45 | **23.77** | 465.56 | | **272.78** |
| D15 | 0.50 | 0.55 | **0.92** | 26.47 | 21.08 | **26.85** | 506.09 | | **134.15** |
| D26 | 0.49 | 0.47 | **0.43** | 30.32 | 20.92 | **25.99** | 517.74 | | **163.66** |
| D50 | 0.48 | 0.58 | **0.95** | 21.61 | 19.13 | **27.34** | 793.91 | | **119.88** |



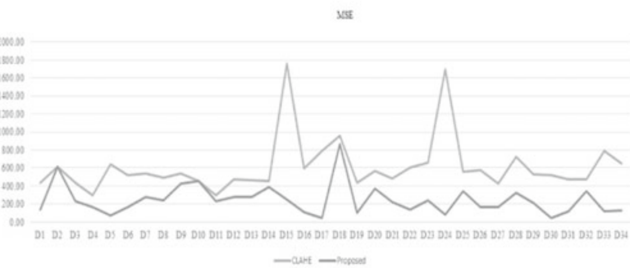

(a) SSIM




(b) PSNR




(c) UIQI




(d) MSE

**Fig. 8** **a** SSIM. **b** PSNR. **c** UIQI. **d** MSE

## 5 Conclusion

A novel technique for enhancement of fused CT and MRI images is proposed. CLAHE is used to enhance the contrast of the fused image and the selection of operational parameters like clip limit, block size and distribution function is made suitably. The block size is chosen for the experiment as $8 \times 8$, with uniform distribution as the transformation function for the enhancement process. The clip limit is in range of 0–0.01, adaptively chosen by PSO algorithm for a maximum fitness value. The proposed technique improves the visual quality. Further, tumors can be segmented and classified efficiently. Experiments show superior results for PSNR, SSIM, UIQI and MSE. This technique can be applied for enhancement of CT-PET images and MRI-PET images. There is a very slight improvement in contrast and entropy for fused image. However, significant change in PSNR, SSIM and UIQI is seen with the proposed enhancement method with reduced MSE. The proposed technique improves the visual quality of the fused images and performs better than CLAHE.

## References


1. Gaillard F et al. Glioblastoma. Article. https://radiopaedia.org/articles/glioblastoma
2. Villanueva JE, Mabray MC, Cha S (2017) Current clinical brain tumor imaging. Neurosurgery 81:397–415
3. Keunen O, Taxt T et al (2014) Multimodal imaging of Gliomas in context of evolving cellular and molecular therapies. Adv Drug Deliv Rev 76:98–115
4. Heiss W-D, Raab P, Lanferman H (2011) Multimodality assessment of brain tumors and tumor recurrence. J Nucl Med 52:1585–1600
5. Verduin M, Compter I, Seijvers D et al (2018) Noninvasive Glioblastoma testing: multimodal approach to monitoring and predicting treatment response. Hindawi Disease Markers 2018. Article ID 2908609
6. Fei Y, Wei G, Zongxi S (2017) Medical image fusion based on feature extraction and sparse representation. Hindawi Int J Biomed Imag 2017
7. Huang H, Feng X, Jiang J (2017) Medical image fusion algorithm based on nonlinear approximation of contourlet transform and regional features. Hindwai J Electr Comput Eng 2017
8. Mzoughi H, Njeh I et al (2018) Histogram equalization-based techniques for contrast enhancement of MRI brain glioma tumor images: comparative study. In: Proceedings of 4th international conference on advanced technologies for signal and image processing (ATSIP), 21–24 March 2018
9. Yang B, Li S (2010) Multifocus image fusion and restoration with sparse restoration. IEEE Trans Instrum Meas 59(4)
10. Fiete RD Modelling the imaging chain of digital camera, image enhancement processing, Chap. 9 [Online]. SPIE Digital Library
11. Chang Y, Jung C et al (2010) Automatic contrast-limited adaptive histogram equalization with dual gamma correction. IEEE Access 6
12. Min BS, Lim DK et al (2013) A novel method of determining parameters of CLAHE based on image entropy. Int J Softw Eng Appl 7(5):113–120

13. Agarwal M, Mahajan R (2017) Medical image contrast enhancement using range limited weighted histogram equalization. In: Proceedings of 6th international conference on smart computing and communication, ICSCC, December 2017
14. Agarwal M, Mahajan R (2017) Medical image contrast enhancement using quad weighted histogram equalization with adaptive Gama correction and homomorphic filtering. In: Proceedings of 7th international conference on advances in computing and communications, ICACC-2017
15. Zhu Y, Huang C (2012) An adaptive histogram equalization algorithm on the image gray level mapping. In: Proceedings of 2012 international conference on solid state devices and materials science. Phys Procedia 25:621–628
16. Zuierveld K (1994) Contrast limited adaptive histogram equalization. In: Graphics gems IV, San Diego, CA, USA. Academic Press Professional, Inc, pp 474–485
17. Kaur H, Rani J (2016) MRI brain image enhancement using Histogram equalization techniques. In: Proceedings of IEEE WISPNET
18. Koonsanit K, Thongvigitmanee S et al (2017) Image enhancement on digital X-ray images using N-CLAHE. In: Proceedings of IEEE 10th biomedical engineering international conference (BMEiCON), August 2017
19. Mohan S, Mahesh TR (2013) Particle swarm optimization based contrast limited enhancement for mammogram images. In: Proceedings of 7th international conference on intelligent systems and control
20. Bhat M, Tarun Patil MS (2014) Adaptive clip limit for contrast limited adaptive Histogram equalization (CLAHE) of medical images using least mean square algorithm. In: Proceedings of international conference on advance communication control and computing technologies (ICACCCT). IEEE
21. More LG, Brizuela MA et al (2015) Parameter tuning of CLAHE-based on multi-objective optimization to achieve different contrast levels in medical images. In: Proceedings of international conference on image processing (ICIP). IEEE
22. Jennifer S, Parasuraman S, Kadirvelu A (2016) Contrast enhancement and brightness preserving of digital mammograms using fuzzy clipped contrast-limited adaptive histogram equalization algorithm. Appl Soft Comput 42:167–177
23. Joseph J, Sivaraman J, Periyasamy R, Simi VR (2017) An objective method to identify optimum clip-limit and histogram specification of contrast limited adaptive histogram equalization for MR Images. Nalecz Inst Biocybern Biomed Eng 37:489–497
24. Singh G, Mittal AK (2014) Controlled bilateral filter and CLAHE based approach for image enhancement. Int J Eng Comput Sci 3(11)
25. Makandar A, Halalli B (2015) Breast cancer image enhancement using median filter and CLAHE. Proc Int J Sci Eng Res 6(4)
26. Zeki M, Erturk S (2015) Enhancement of ultrasound images with bilateral filter and Rayleigh CLAHE. In: Proceedings of 23rd signal processing and communications applications conference (SIU), Malatya, Turkey
27. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of international conference on neural networks, vol 4. IEEE, pp 1942–1948
28. Braik M, Sheta A, Ayesh A (2007) Image enhancement using particle swarm optimization. In: Proceedings of world congress on engineering, vol 1, London, UK
29. Chandrashekar L, Sreedevi A (2017) Assessment of non-linear filters for MRI Images. In: Proceedings of second IEEE international conference on electrical, computer and communication technologies, 22–24 February 2017, Coimbatore, India
30. Chandrashekar L, Sreedevi A (2018) A hybrid multimodal medical image fusion technique for CT and MRI brain images. IGI Glob Int J Comput Vis Image Process (IJCVIP) 8(3)
31. Chandrashekar L, Sreedevi A (2017) A novel technique for fusing multimodal and multiresolution brain images. Procedia Comput Sci 115:514–548
32. Chandrashekar L, Sreedevi A (2020) A multi objective enhancement technique for poor contrast magnetic resonance images of brain Glioblastomas. Procedia Comput Sci 171:1770–1779

33. Chandrashekar L, Sreedevi A (2020) A two-stage multi-objective enhancement for fused magnetic resonance image and computed tomography brain images. J Inform Syst Telecommun (JIST) 8(2)

# Recent Advances and Future Directions of Assistive Technologies for Alzheimer's Patients

**V. Mohan Gowda and Megha P. Arakeri**

## 1 Introduction

According to Alzheimer's worldwide statistics-2019, nearly more than 50 million people are suffering from Alzheimer's or dementia related diseases around the world out of which 4.1 million in India alone. The number of peoples having Alzheimer's disease is growing faster all over the world. One out of nine people age 65 and older have this disease. Nearly half of the dementia people across the globe live in Asia. Alzheimer's Association estimated the cost of Alzheimer's and dementia related diseases at about \$605 billion and it is about 1% of the gross domestic product of the world [1]. Dementia is the incapability and dependency of elderly people on caretakers. It significantly decreases the mental performance and one of the considerable types of brain disease. Therefore these patients depend more on caretakers. Its symptoms start with a progressive loss of memory, thinking, and analysis. There are different causes behind this process such as blood rushes into the brain, stroke, malnutrition or traumatic conditions. The most common types of dementia are as follows: Alzheimer's disease (AD), vascular dementia, dementia with Lewy bodies, mixed dementia, Parkinson's disease, front temporal lobar degeneration, Creutzfeldt–Jakob disease and normal pressure hydrocephalus [2].

Alzheimer's disease (AD) is the most common form of brain disorder among older people aged above 60. Dementia is a disorder of the brain that seriously affects the ability of a person to carry out routine activities. Firstly, AD affects major parts of

V. Mohan Gowda
Department of CSE, GITAM School of Technology, Bengaluru, India
e-mail: mgowda@gitam.edu

M. P. Arakeri (✉)
Department of ISE, Center of Imaging Technologies, Ramaiah Institute of Technology, Bengaluru, India
e-mail: meghaparakeri@msrit.edu

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_3

the brain which manage thinking, memory and language control. Later people suffer from difficulty remembering conversations, names or events and are which an early clinical symptom [1]. Apathy and depression are also often observed as signs of early symptoms. A later symptom consists of impaired communication, disorientation, confusion, poor judgment capacity, changes in behavior and ultimately difficulty speaking, swallowing and walking. They may forget their daily routine of brushing their teeth and combing hair. There are seven stages of the Alzheimer's disease Stage 1: Normal outward behavior, Stage 2: Very mild changes, Stage 3: Mild decline, Stage 4: Moderate decline, Stage 5: Moderately severe decline, Stage 6: Severe decline and Stage 7: Very severe decline [3].

New emerging technologies in science and engineering are being explored to support individuals with decreased cognitive abilities resulting from Alzheimer's patients. Such technologies range from low technology aids to higher technology aids [4]. Dew [5] describes Intelligent Assistive Technologies (IAT) is one such technological tool that has been developed to reduce the burden of Alzheimer's patients and their caretakers. IAT has been created to compensate for the loss of memory, executive function and offer remarkable assistance to older adults with dementia or Alzheimer's in discharging their day to day activities. A wide verity of IAT has been developed. People suffering from Alzheimer's decease and their caretakers would be very much relieved if these technological tools assist them in performing their daily activities like bathing, cooking, hand wash, brushing and medication aids etc. Further, a safety system that ensures the safety of the patients in the event of a fall, gas leakage detection and fire etc. is also essential. In this paper, the various existing IAT for medication aids, brushing and fall detection are visited and a summary of the recent developments, advancements, merits and demerits of each IAT techniques are explored.

## 2 Assistive Technology (AT) as an Assisting Aid for Alzheimer's Patients

Assistive technology is an umbrella term that forms the basis for the working of devices that assist, adapt and rehabilitate the people with disabilities or the aged population with reduced cognitive functions. Similar to most of the mobility aids like walkers, scooters, wheelchairs, crutches, prosthetics devices and podiatry arch supports. IAT also helps compensate for certain impairments and promotes a person's feeling of being independent decreasing the need or the other's support in discharging his day-to-day activity. The caregivers are using different types of assistive technology like safety (Baby monitors, GPS trackers to manage wandering, Motion sensors to monitor patient movement, etc.), cognitive engagement and memory devices (Talking tiles to record reminders, Talking and atomic blocks and etc.) and health (Fitness tracking app, Automatic locking pill dispensers and etc.) [6]. Thus, IAT is a cutting edge technology geared for Alzheimer's care and the technology for Alzheimer's deceases falls into four categories [7]. (a) Caregiver peace of mind (b) Caregivers education, organization, and planning (c) Patient assistance and (d) Patient enrichment.

Figure 1 shows the taxonomy table of several strategies and technologies of AT aimed at providing intelligent tools such as Memory aids (Memory glasses[8], Memory chip[9], Planning and Execution Assistant and Training System (PEAT) [10], Remember alert (RA) [11]), daily activities (COACH [12], VERA [5], cook's collage) and safety activities (CareWatch [13], CareMedia [14] and Fall detector [15]). These virtual assistants can be great for people with Alzheimer's not only in terms of helping them with daily tasks but also as reminders and so on. Coming to medication aid technologies, there are automatic dispensers that can alert the patients to take medicine. These systems will also alert the caregiver whether the medication is done or not. However, these assistive technologies will not provide step by step process to the patients to be followed in taking the medicine. Even though some tools aid the patient in hand washing, no accurate assistive technologies for other daily activities like brushing and bathing have been reported to the best of our knowledge. Another assistive technology is to provide the safety mechanism which should detect some of the activities like fall detection. Though there are some systems which can detect a fall both manually and simple alarm automated system exist to notify the

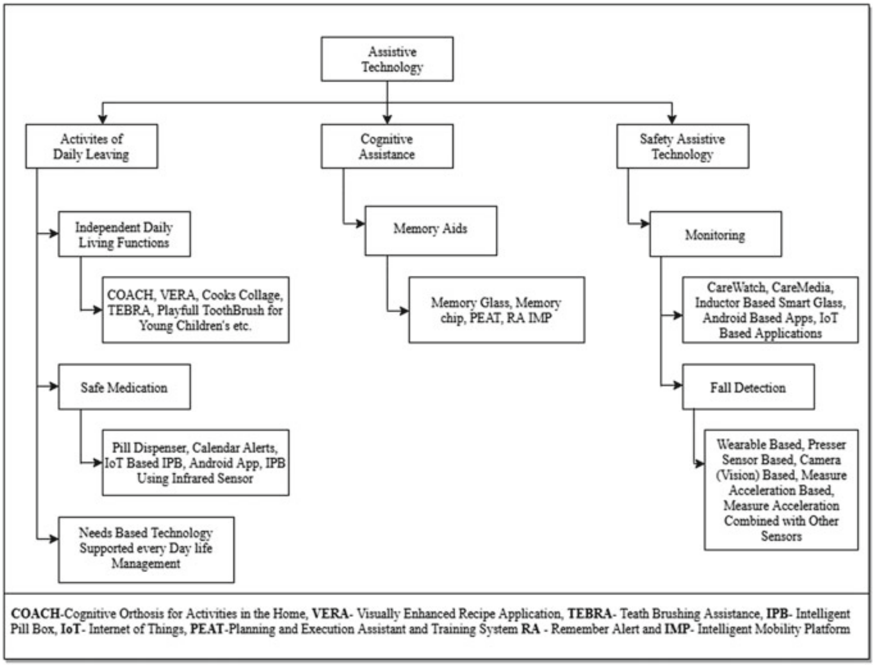



**Fig. 1** Taxonomy table of assistive technology

caregivers of a fall. The automated devices must be worn by the patient 24/7 and the manual ones require the user to activate the response system in the event of the fall.

## 3 Assistive Technologies for the Detection of Fall

Various techniques have been devised to detect various possible falls that may occur during walking, standing, even climbing up or down from ladders, etc. Major steps in the automated detection of fall and alert system are shown in Fig. 2. The classification of fall detection and daily activity is not an easy task as daily activities like sitting down or going to stand position to lying down are similar to fall. Based on the type of technology used, like wearable devices, pressure sensor and camera (vision), several fall detection methods are introduced in the following section.

### *3.1 Fall Detection Methods Based on Wearable Devices*

Fall detection methods based on wearable devices are widely used in both indoor and outdoor owing to their low cost and undisturbed in terms of privacy. Erdogan [16] proposed the wireless sensor network method to detect a fall. First, it collects continuous data from the sensor devises and detects fall occurrences and classify data using the k-nearest neighbour method that measures the similarity and dissimilarity between the instances based on distance function. However, this method is not detecting a different kind of fall and it also doesn't inform the caretaker or family members in a short time. Diep [17] proposed the feature extracted schema for the continuous data stream of the sensors using a super vector machine. It classifies the falls or not falls. These methods do not automatically calculate the bin and range values. Putra [18] developed an event-triggered machine learning approach with multipeak detection using an accelerometer sensor to detect falls. This method uses to align fall stages by finite state machines based on feature extraction with a segment. It also resolves the ambiguity problem in multiple pick detection. But this system's high computational time leads to the consumption of more battery. The fall

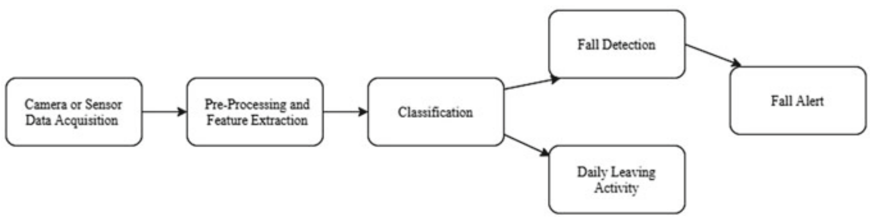



**Fig. 2** General automated fall detection and alert system

detection rate is also less because it takes data only from the accelerometer sensor. To overcome this issue, Sucerquia [19] developed a periodic fall detection method using nonlinear classification and Kalman filter to reduce the false result. This system is less energy efficient because it requires only a 25 Hz sampling rate. This also doesn't require a large memory. However, patients must wear a device near the waist only otherwise difficult to identify the fall. Also, the periodic detector can work only if the device is upside-down otherwise it doesn't work properly. Khojasteh [20] individually collected the data of elder and compared several methods to improve the fall detection rate. In Nait [21] high-risk fall detection is compared with deep learning models which are convolution neural network, long short-term memory (LSTM) and a combination of these two (ConvLSTM) to each other. Data was collected through an accelerometer sensor and it matched with an existing dataset of 296 adults. However, it tests the existing dataset of only 256 adults. Santos [22] implemented the Internet of things (IoT) and fog computing environment based deep learning fall detection. They proposed a convolution neural network to extract relevant features from the data collected by the accelerometer sensor. It resolves the issues of dimensionality, simplicity and improves the performance of the training data using data augmentation. But the convolution neural network is of high computational cost. Mauldin [23] developed a smartwatch IoT device system using deep learning to detect fall and data collection by an accelerometer sensor with the help of an android app. The wearable sensor device is used to carry every ware and the cost of the sensor is also low. They evaluated with supper vector machine, naive bayes and deep learning algorithms to detect a fall using different fall datasets. However, they have not been evaluated with real word patients. Some commercial wearable devices are also available such as Apple Watches [24] and Plilips Life Line Auto Alert [25].

### *3.2 Fall Detection Methods Based on Pressure Sensor*

Fall detection based wearable's methods has the main drawback of elder people having inconvenience in wearing the device all the time. Suppose if they forget to wear a device, then it is difficult to notify the fall. To overcome this problem floor sensor-based fall detection methods were introduced. In this approach, wearing any device is not required by the patient. Rimminen [26, 27] developed a method of tracking people using a near field image (NFI) floor sensor. They used a human tracking system to sense human presence using the electric near field. This improves the positioning performance of moving target using unprocessed observation, observation centroids and Kalman filtered centroids. Feng [28] proposed a floor pressure imaging system to detect fall using smart floor fixed with fiber optic sensors. Fall is detected based on lying on the floor for more than the threshold value. If an individual is lying on the floor for more than the threshold values, a fall is detected. The drawback in this method is pressure based floor sensor cannot differentiate between an actual fall and an individual lying on the floor more than the threshold value. Dahe [29] merged the force sensor and accelerometer sensor under intelligent tiles to detect the fall of the individually living elders. This method improves the performance of fall detection. Another drawback of pressure sensor based methods is difficulty in installation and maintenance.

### 3.3 Fall Detection Methods Based on Vision

Wearable based fall detection needs to wear devices all the time. Pressure Sensor based fall detection also gives false result when the patient individually lying on the floor for more duration. Thus, vision based fall detection methods were recommended and are useful for elder people to identify falls in the restricted area. Zhong [30] Vision-based methods are categorized into 3 types. (a) Fall detection using a single RGB camera, (b) 3D-based methods using multiple cameras, (c) 3D-based methods using depth cameras. In single RGB camera methods, it is difficult to capture different kinds of surfaces and it capture only limited distance. 3D-based methods using multiple cameras that achieve better results on different kinds of surfaces but need to improve the reliability of tracking. Depth-camera-based methods usually perform better reliability of tracking since the depth camera is not affected by the changes in lighting. Planinc [31] introduced fall detection by a combination of 3D tracking data obtained by the Kinect with fuzzy logic. Using 3D reconstruction by the Kinect, it works during day and night time also. Hence the accuracy is increased. Alazrai [32] proposed Motion-Pose Geometric Descriptor (MPGD) to build an invariant description for human activity using Kinect sensors. From the video frames extracted within each segment, they constructed a histogram based representation (HBR) to the MPGD and a super vector machine was trained to predict the probabilistic output of the activities performed in the partially observed video. Using the Kinect sensor, they simulated the fall and non-fall related activity of elderly people like walking, sitting, falling from standing, and falling from sitting. MGPD uses a single Kinect sensor, hence it covers only a limited distance and subject occlusion problem. Yao [33] in this work human torso motion model (HTMM) was proposed that considers several points between the shoulder center and hip center. HTMM considers 20 points and 10 points on the human body while a person is standing and sitting positions respectively. One of the major limitations of this technique is it does not produce correct results for vision captured at a distance of beyond 7 m. Nizam [34] proposed a fall detection algorithm to identify the fall risk level at the current time before fall. The accuracy of fall detection is improved. But results mainly depend on the skeleton detection of the Microsoft Kinect sensor value. However, it was evaluated only on the health peoples. Miguel [35] proposed artificial vision-based home automation using a machine learning decision system to detect falls. But they need to improve the performance during occlusion and change the lighting.

**Table 1** Confusion matrix for fall detection

| | Fall truth identified (Positive) | Fall truth not identified (Negative) |
|---|---|---|
| Fall images present in the dataset | True positive (TP) | False negative (FN) |
| Fall images not present in the dataset | False positive (FP) | True negative (TN) |

True positive = the number of fall images that are identified correctly
True negative = the fall images that are not present in the dataset and not identified
False positive = the number of fall images which are not matching but identified as one of the falls in the database
False negative = the number of falls that are of the database but not identified correctly

## 3.4 *Performance Evaluation Parameters*

Various performance parameters for evaluating fall detection methods are discussed below [36]. Table 1 shows the Confusion matrix for fall detection.

- **Precision** is what percentage of fall that the classifier labeled as positive is actually positive.

$$\text{Precision} = \text{TP}/(\text{TP} + \text{TN})$$

- **Recall** is what percentage of positive fall the classifier labeled as positive.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN})$$

- **Accuracy** gives the percentage of data set falls that are correctly classified.

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

- **Sensitivity** is true positive recognition rate.

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN})$$

- **F Score** is a weighted average of the recall and precision.

  Table 2 summarizes the various fall detection methods.

## 3.5 *Existing Datasets*

The different datasets for different types fall (falling down, bending, squatting, sitting, lying and walking, etc.) to evaluate fall detection algorithms. These datasets are publicly available listed in Table 3.

**Table 2** Summary of fall detection techniques

| Technique | Sensor/device | Performance | Reference |
|---|---|---|---|
| Pose invariant Random Decision Tree (RDT) algorithm, SVM classifier | RGB Camera | Accuracy—97.6% Sensitivity—95.3% | Bian [37] |
| Wireless sensor network, K-nearest neighbour | Dual-Axis Accelerometer, Dual-Axis Magnetometer | Accuracy—89.4% | Erdogan [16] |
| Feature extraction, SVM classification | Low cost wearable sensor | Precision—91.9% Recall—94.4% | Diep [17] |
| Event-triggered machine learning approach | Accelerometer | Cogent dataset Precision—90.9% Recall—87.6% SisFall dataset Precision—54.1% Recall—100% | Putra [18] |
| Nonlinear classification Kalman filter | Triaxial Accelerometer | Accuracy—99.4% | Sucerquia [19] |
| Convolution neural network, IoT and fog computation | Accelerometer | Accuracy—99.68% | Santos [21] |
| The machine learning algorithm, IoT application | Smartwatch Accelerometer | Accuracy—86% | Mauldin [23] |
| Feature extraction, pose estimation step | Floor sensor | Sensitivity—91% | Rimminen [26] |
| Reel-based triangulation system, Kalman filter | Floor sensor | Accuracy: when the distance between the two people is over 0.8 m is 90% and1.1 m is 99% | Rimminen [27] |
| Posture-based fall detection | Fiber sensor | – | Feng [28] |
| Intelligent tiles | Force sensor, 3-axis Accelerometer | Sensitivity >91% | Daher [29] |
| Feature extraction, fuzzy logic for robust | 3D camera, Kinect | Accuracy—98.6% | Planinc [31] |
| Human torso motion model (HTMM) | Kinect, 3D Depth camera | Accuracy—97.5% | Yao [25] |
| Motion pose geometric description (MPGD), Histogram based representation (HBR), Multi-class SVM classifier | Microsoft Kinect sensor, Depth camera | Accuracy—For fully-Observed Video Sequences-93.6%. For single and two unobserved video subsequences with random length-77.6% and 65% respectively | Alazrai [32] |

(continued)

**Table 2** (continued)

| Technique | Sensor/device | Performance | Reference |
|---|---|---|---|
| Data acquisition and generation of skeleton data, Identification of fall risk factors, SVM classification | Microsoft Kinect sensor, Depth camera | Accuracy—88.57%<br>Sensitivity—96.67%<br>Precision—80.56% | Nizam [34] |
| Smart home based on artificial vision algorithm, KNN classification | Camera, Raspberry Pi, Wifi adapter | Accuracy—96% | DeMiguel [35] |

## 4 Assistive Technology for Monitoring Daily Activities

The dementia people require full assistance from a caregiver for completing activities of daily living (ADL) like hand wash, bathing, dressing, brushing, medication aids and cooking so on. The following section focuses on assistive technologies for medication aid and brushing.

### *4.1 Intelligent Medication System*

The elderly and Alzheimer's patients often suffer from memory loss. This problem indirectly reflects on the health issues of these people because they lack the commitment to take medicine. Figure 3 shows the intelligent medication system. Initially, pill timing in the system needs to be set according to the patient's requirement. The system will trigger an alarm at that fixed time. By adding an LCD screen which shows the time, the system can be made user friendly. Also, the system automatically sends a notification to the caretaker about medicine taken. Mayuresh [43] proposed the Ardumed-smart medicine system gives reminder messages to patients in three different ways. First, it gives a light indication. If a patient is not close to the system then he may not notify light indication. Hence after some time, it gives the sound notification. Suppose the patient is outside, then they could not able to notify the sound also. Thus finally it sends the mobile notification using the android app. Hussain [44] proposed mobile application remembering them date/time and amount of medicine to consume. Eric [45] developed a Radio frequency identification (RFID) based smart drawer for monitoring and recording of the patient taken medicine or not. It also helps the caregiver to retrieve historical data of the patient taking medicine. Shashank [46] developed an IoT based smart pillbox with reminding and consumption. The proposed pillbox is given alert sound to the user for a particular time to take pills and required pills automatically come out of the smart pillbox to avoid confusion among medicine. Shih-Chang [47] An Intelligent Pill Box (IPB) was designed. The IPB can send the medicine bag out of the box and remind the elder/patients to take medicine in time. Suppose a patient does not take the medicine bag out of the box, then

**Table 3** Summary of fall detection datasets

| Dataset | Size | Purpose |
|---|---|---|
| UFRD [34] | Videos of falls are 30 and videos of Activities of Daily Living (ADL) are 40. Total frames are11,936 | Vision based Benchmark dataset to evaluate unconstrained fall detection |
| Multicam [38] | It contains 24 performances and 261,137 frames. Each performance has been recorded from 8 different views | To test the simulated fall and normal daily activities acquired in realistic studies |
| FDD [39] | 108,476 frames contain 4 different stages with multiple actors | The robustness of the method is evaluated for location change between training and testing |
| SDUFall [30] | One Kinect has been installed to capture the ten young men and women did six types of activities 30 times, finally 1800 video clips | To study the impact of light effect, direction and position changes concerning the camera |
| European Data Format [30] | Two Kinect have been installed to capture the events from two viewpoints, leading to a total of 320 sequences of five different ADL (subject) | Classification of patient action from a specific viewpoint |
| OCCU [30] | The dataset contains occluded falls of the same subject from two different viewpoints<br>The first viewpoint contains 25,618 frames and 30 occluded falls<br>The second viewpoint contains 23,703 frames and 30 occluded falls | Using different subjects and viewpoints, user independence and viewpoints independence are evaluated |
| UP dataset [40] | This contains 40 videos of people doing normal ADL and 30 videos with a fall in them. These videos are more than 850 GB | Helps Research communities of human activity recognition and machine learning to compare their fall detection solutions fairly |
| TST V2 [41] | The dataset is composed of ADL and fall simulated by 11 young actors to generate 264 sequences. It contains 46,418 frames | It is used for depth frames and skeleton joint scenarios |
| SIS fall [42] | It consists of 2706 ADL events and 1798 falls events | Used for wearable fall detection. It has a uniform length (15 s) for its fall data that gives a better F-score for the classifier |

(continued)

**Table 3** (continued)

| Dataset | Size | Purpose |
|---|---|---|
| Cogent [18] | It has 644 fall and 1196 ADL events | Used for wearable fall detection. It has a dynamic length for its fall data. Hence it is easy to predict the length of human activity |

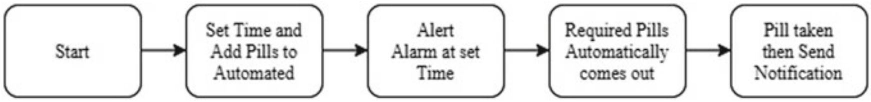



**Fig. 3** General automated medication aid

IPB remotely informs the family members via skype. So they can help the patients. But IPB works only if the internet connection is available. Jayesh [48] developed an autonomous pill dispenser box containing a programmable alarm system with an interactive user interface that sends a notification about medicine taken and it also sends a message to the medical supplier if the pills are less in the pill dispenser box. Juan [49] developed IoT based intelligent pillbox is a programmable alarm system with automatic opening and closing of the box at a particular time and it also gives the notification message to the caregiver using Global System for Mobile (GSM) module. Table 4 shows the medication system's key techniques with its results and Issues.

### *4.2 Intelligent Brushing System*

The dementia peoples face difficult or impossible to brush their teeth on their own. By using some assistive technologies to overcome these difficulties, Chang [50] developed a playful toothbrush for young children. The system is vision-based and recognizes different toothbrushing motions. A fun tooth brushing game is displayed on an LCD which takes it as an input to a child's physical brushing his/her teeth. Flagg [51] designed an automated and intelligent promoting application vision based toothbrush tracker for brushing teeth. The vision based toothbrush tracker motion and interaction facial features of the human behavior, toothbrush, toothpaste, etc. intelligent promoting application is created using Partially Observable Markov Decision Process. Peters [52] developed the TEBRA system to support mildly impaired people during teeth-brushing. A user study was performed with 7 participants suffering from cognitive deficits. The results showed that users made significantly more independent steps when they had access to the system's prompts.

**Table 4** Summary of medication aids

| Technique | Result | Issues | Reference |
|---|---|---|---|
| Ardumed-smart medicine system using Android App | Gives alert message to 3 different ways<br>Vision indication<br>Sound alarm<br>Mobile notification | The Ardumed-smart medicine system is not automated | Mayuresh [43] |
| Mobile reminder system for Alzheimer's patients | Alert date/time and amount of medicine to consume | The Alert system is not automated | Abu-Dalbouh [44] |
| RFID based smart drawer | The smart drawer is monitoring and recording medicine is taking or not from the drawer | No alert system is not fully automated | Eric [45] |
| IoT based smart pillbox | Alert the patients to take pills, the required pills automatically come out of the box | Portability issues, application compatibility issues | Shashank [46] |
| Intelligent Pill Box (IPB) using an Infrared sensor | IPB is an automated dispensary system if the patient is not taking medicine bag out of the box. IPB informed the caregiver remotely via Skype | IPB works only if the internet is available | Shih-Chang [47] |
| Autonomous pill dispenser Box using IoT | The system contains a programmable alarm system with an interactive UI. It sends a notification to caregiver medicine taken and also send a message to medicine suppliers if he medicine is less in the pillbox | Portability issue, It works only on the internet | Jayesh [48] |
| IoT based intelligent pillbox | Programmable alarm system with automated open and close the pillbox. It also gives the notification message to the caregiver using the GSM module | Portability issues, network disruption Issues | Juan [49] |

## 5 Challenges

Detection of falls and assistance to daily activities of Alzheimer's patients is not an easy task. Most fall detection methods are based on wearable devices, pressure sensors and vision. Daily activities such as brushing and medication aid are based on sensors and IoT applications. This section emphasizes the challenges involved in fall detection and daily activities to Alzheimer's patients and also gives an outline of future research work that can be carried out in this field.

- *Recognition of fall under varied illumination and low resolution:* This approach has few drawbacks due to the following reasons: (1) Poor quality of images captured in cameras include a large distance between the subject and the camera, (2) Reduced spatial resolution of the camera, (3) Speed at which the subject is moving, (4) Changes in the illumination at the monitored location.
- *Increase the computational costs:* The complexity increases generally with the size of the input. However, the complexity of an algorithm may vary dramatically for different inputs of the same size.
- *Fall detection under partial occlusion:* As the subject is not anticipated to be Cooperative, it may be blocked by other moving objects or the subject may cover his face from the camera deliberately not to reveal his identity.
- The complexity of the fall detection algorithm, not all fall situations are detected with the same certainty, as a wearable device, old people tend to forget wearing it.
- Another challenge is to wearable based fall detection is that patient has to wear devices all the time. Suppose if forget to wear a device, it is difficult to monitor.
- Pressure based floor sensors cannot differentiate between fall and individual laying on the floor more than the threshold time. Also, it is difficult for installations and maintenance.
- Lack of step by step detailed procedures to guide Alzheimer's patients to perform medication aids and brushing activities incorrect manner.
- Guides the activity for Alzheimer's people can complete, independently from a caregiver.

## 6 Future Directions

All the technologies discussed above were aimed at providing an intelligent system for fall detection, Medication aid and brushing.

- Fall detection is based on a wearable device, pressure sensor and vision.
  - Wearable based fall detection methods are widely used in both indoor and outdoor owing to their low cost and undisturbed in terms of privacy. But wearable based fall detection techniques require that patients must have to wear devices all the time otherwise it is difficult to monitor.

  - Pressure Sensor based fall detection scheme gives false results when a patient individually laying on the floor for a long time. Therefore, vision based fall detection methods are useful for elder people to identify falls in the restricted area.
  - Computer vision based fall detection models also have the following drawbacks.
    i False positive indicates the number of fall images which are not matching but identified as one of the falls in the database.
    ii False negative indicates the number of falls but not falling, which is voluntarily bent down. Due to false positives and false negative results, computer vision based systems/algorithms do not produce accurate results in most of the scenarios.

- Medication aid technologies employ automatic dispensers which alert patients to take medicine at the prescribed time. However, these assistive technologies do not provide step by step process on how to take medicine to Alzheimer's patients.
- The exit brushing system developed people with moderate cognitive disabilities and young Childs. But they do not provide step by step process to complete the task.

Based on the above discussion and review, we propose to develop self-assistive technology that assists the people who are suffering from Alzheimer's disease in activities such as brushing, medication aides and fall detection with less or zero dependency on caretakers.

## 7 Conclusion

IATs have a profound impact in the medical field particularly for the people suffering from Alzheimer's patients. We believe that the paper has given a brief and systematic review account of safety assistive tools about fall detection, daily activity tools on brushing and medication. We have reviewed various assistive technology systems concerning their working principle, advantages and limitations for fall detection. It is also presented a systematic review of systems for Medication aids and brushing, ranging from a simple alarm system to automated IoT based systems. This paper also presented available standard databases and performance evaluation techniques for fall detection. We strongly believe that this paper paves the way to other researchers working in this area to know the past and future challenges.

## References

1. Alzheimer's disease facts and figures. https://www.alz.org/media/documents/alzheimers-facts-and-figures-2019

2. Klimova B, Maresova P, Kuca K (2016) Assistive technologies for managing language disorders in dementia. Neuropsychiatr Dis Treat 12:533. https://doi.org/10.2147/NDT.S95903
3. Alzheimer's disease stages. https://www.alzheimers.net/stages-of-alzheimers-disease/
4. Ienca M, Jotterand F (2016) Intelligent assistive technologies for dementia: social, legal and ethical implications
5. Dew MA, Bharucha J, Anand V, Forlizzi J, Reynolds CF III, Stevens S, Wactlar H (2009) Intelligent assistive technology applicationsto dementia care: current capabilities, limitations, and future challenges. Am J Geriatr Psychiatry 17:88–104. https://doi.org/10.1097/JGP.0b013e318187dde5
6. Technological innovations for those with dementia. https://www.alzheimers.net/9-22-14-technology-for-dementia/
7. Assistive technology for Alzheimer's. https://togetherinthis.com/what-is-assistive-technology-for-alzheimers/
8. DeVaul RW (2004) The memory glasses: wearable computing for just-intime memory support. Ph.D. dissertation, Massachusetts Institute of Technology
9. Beigl M (2000) Memoclip: a location-based remembrance appliance. Pers Technol 4:230–233. https://doi.org/10.1007/BF02391564
10. Levinson R (1997) The planning and execution assistant and training system. J Head Trauma Rehabil 12:769–775
11. Using technology to help with everyday life. https://www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=109
12. Mihailidis, Boger JN, Craig T, Hoey J (2008) The coach prompting system to assist older adults with dementia through handwashing: an efficacy study. BMC Geriatr 8:28. https://doi.org/10.1186/1471-2318-8-28
13. Rowe M, Lane S, Phipps C (2007) Carewatch: a home monitoring system for use in homes of persons with cognitive impairment. Topics Geriatr Rehabil 23:3–8
14. Bharucha, Atkeson C, Chen D (2006) Caremedia: automated video and sensor analysis for geriatric care. In: Annual meeting of the american association for geriatric psychiatry
15. Torbay council: fall detector, 2007. Available at http://www.torbay.gov.uk/index/socialcare/help-in-your-home/community-alarm-services/fall-detector.html
16. Erdogan SZ, Bilgin TT (2012) A data mining approach for fall detection by using k-nearest neighbour algorithm on wireless sensor network data. IET Commun 6:3281–3287. https://doi.org/10.1049/iet-com.2011.0228
17. Diep NN, Pham C, Phuong TM (2013) A classifier based approach to real-time fall detection using low-cost wearable sensors. In: 2013 international conference on soft computing and pattern recognition (SoCPaR). IEEE, pp 105–110. https://doi.org/10.1109/SOCPAR.2013.7054110
18. Putra I, Brusey J, Gaura E, Vesilo R (2018) An event-triggered machine learning approach for accelerometer-based fall detection. Sensors 18:20. https://doi.org/10.3390/s18010020
19. Sucerquia, Lopez JD, Vargas-Bonilla JF (2018) Real-life/real-time elderly fall detection with a triaxial accelerometer. Sensors 18:1101. https://doi.org/10.3390/s18041101
20. Khojasteh SB, Villar JR, Chira C, González VM, De la Cal E (2018) Improving fall detection using an on-wrist wearable accelerometer. Sensors 18:1350. https://doi.org/10.3390/s18051350
21. Aicha N, Englebienne G, Van Schooten KS, Pijnappels M, Kröse B (2018) Deep learning to predict falls in older adults based on daily-life trunk accelerometry. Sensors 18:1654. https://doi.org/10.3390/s18051654
22. Santos GL, Endo PT, Monteiro KH, Rocha ED, Silva I, Lynn T (2019) Accelerometer-based human fall detection using convolutional neural networks. Sensors 19:1644. https://doi.org/10.3390/s19071644
23. Mauldin TR, Canby ME, Metsis V, Ngu AH, Rivera C (2018) Smartfall: a smartwatch-based fall detection system using deep learning. Sensors 18:3363. https://doi.org/10.3390/s18103363
24. AppleWatch. Available online https://support.apple.com/en-us/HT208944. Accessed on 21 July 2019

25. Lifeline P. Available online https://www.lifeline.philips.com/medical-alert-systems/homesafeautoalert.html. Accessed on 21 July 2019
26. Rimminen H, Lindström J, Linnavuo M, Sepponen R (2010) Detection of falls among the elderly by a floor sensor using the electric near field. IEEE Trans Inform Technol Biomed 14:1475–1476. https://doi.org/10.1109/TITB.2010.2051956
27. Rimminen H, Lindström J, Linnavuo M, Sepponen R (2009) Positioning accuracy and multi-target separation with a human tracking system using near field imaging. Int J Smart Sens Intell Syst 2:156–175
28. Feng G, Mai J, Ban Z, Guo X, Wang G (2016) Floor pressure imaging for fall detection with fiber-optic sensors. IEEE Pervasive Comput 15:40–47. https://doi.org/10.1109/MPRV.2016.27
29. Daher M, Diab A, El Najjar MEB, Khalil MA, Charpillet F (2016) Elder tracking and fall detection system using smart tiles. IEEE Sens J 17:469–479. https://doi.org/10.1109/JSEN.2016.2625099
30. Zhang Z, Conly C, Athitsos V (2015) A survey on vision-based fall detection. In: Proceedings of the 8th ACM international conference on pervasive technologies related to assistive environments, pp 1–7. https://doi.org/10.1145/2769493.2769540
31. Planinc R, Kampel M (2012) Robust fall detection by combining 3D data and fuzzy logic. In: Asian conference on computer vision. Springer, pp 121–132. https://doi.org/10.1007/978-3-642-37484-5_11
32. Alazrai R, Momani M, Daoud MI (2017) Fall detection for elderly from partially observed depth-map video sequences based on view invariant human activity representation. Appl Sci 7:316. https://doi.org/10.3390/app7040316
33. Yao L, Min M, Lu K (2017) A new approach to fall detection based on the human torso motion model. Appl Sci 7:993. https://doi.org/10.3390/app7100993
34. Nizam Y, Mohd MNH, Jamil M (2018) Development of a user adaptable human fall detection based on fall risk levels using depth sensor. Sensors 18:2260. https://doi.org/10.3390/s18072260
35. Miguel KD, Brunete A, Hernando M, Gambao E (2017) Home camera-based fall detection system for the elderly. Sensors 17:2864. https://doi.org/10.3390/s17122864
36. Lakshmi N, Arakeri MP (2018) Face recognition in surveillance video for criminal investigations: a review. In: International conference on communication, networks and computing. Springer, pp 351–364. https://doi.org/10.1007/978-981-13-2372-0_31
37. Bian Z-P, Hou J, Chau L-P, Magnenat-Thalmann N (2014) Fall detection based on body part tracking using a depth camera. IEEE J Biomed Health Inform 19:430–439. https://doi.org/10.1109/JBHI.2014.2319372
38. Ali SF, Khan R, Mahmood A, Hassan MT, Jeon M (2018) Using temporal covariance of motion and geometric features via boosting for human fall detection. Sensors 18:1918. https://doi.org/10.3390/s18061918
39. Fall-detection-Dataset. http://le2i.cnrs.fr/Fall-detection-Dataset?lang=en
40. Seredin O, Kopylov A, Huang S-C, Rodionov D (2019) A skeleton features-based fall detection using microsoft kinect v2 with one class classifier outlier removal. In: International archives of the photo grammetry, remote sensing and spatial information sciences. https://doi.org/10.5194/isprs-archives-XLII-2-W12-189-2019
41. Martíinez-Villaseñor L, Ponce H, Brieva J, Moya-Albor E, Nunez-Martmez J, Penafort-Asturiano C (2019) Up-fall detection dataset: a multimodal approach. Sensors 19:1988. https://doi.org/10.3390/s19091988
42. Sucerquia López JD, Vargas-Bonilla JF (2017) Sisfall: a fall and movement dataset. Sensors 17:198. https://doi.org/10.3390/s17010198
43. Waykole M, Prakash V, Nalini NHS (2016) Ardumed-smart medicine reminder for old people. Int J Sci Eng Res 7:650–654
44. Abu-Dalbouh H, Al-Habeeb A, Al-Kholifi A, Al-Motiry I, Al-Buhairy M (2015) A mobile reminder system for elderly and alzheimer's patients. Int J Comput Sci Issues (IJCSI) 12:95
45. Becker E, Metsis V, Arora R, Vinjumur J, Xu Y, Makedon F (2009) Smartdrawer: Rfid-based smart medicine drawer for assistive environments. In: Proceedings of the 2nd international

conference on pervasive technologies related to assistive environments, pp 1–8. https://doi.org/10.1145/1579114.1579163
46. Shinde S, Kadaskar T, Patil P, Barathe R (2017) A smart pill box with remind and consumption using IoT. Int Res J Eng Technol 4:152–154
47. Huang S-C, Chang H-Y, Jhu Y-C, Chen G-Y (2014) The intelligent pill box7design and implementation. In: 2014 IEEE international conference on consumer electronics, Taiwan. IEEE, pp 235–236. https://doi.org/10.1109/ICCE-TW.2014.6904076
48. Patil J, Khairmode S, Lokhande R, Shinde O (2019) The autonomous pill dispenser with alarm and mobile notifications. Int J Res Appl Sci Eng Technol (IJRASET) 7:2502–2508
49. Parra JM, Valdez W, Guevara A, Cedillo P, Ortiz-Segarra J (2017) Intelligent pillbox: automatic and programmable assistive technology device. In: 2017 13th IASTED international conference on biomedical engineering (BioMed). IEEE, pp 74–81. https://doi.org/10.2316/P.2017.852-051
50. Chang Y-C, Huang C-J, Lo J-L, Chu H-H (2007) A playful toothbrush to motivate proper brushing for young children. In: Demo session and adjunct proceedings of the 9th international conference on ubiquitous computing (UBICOMP 2007)
51. Flagg, Boger J, Mihailidis A (2011) An intelligent toothbrush:machines for smart brushing. In: Abstract presented at the RESNA/ICTA meeting: advancing rehabilitation technologies for an aging society, Toronto, Canada
52. Peters, Hermann T, Wachsmuth S, Hoey J (2014) Automatic task assistance for people with cognitive disabilities in brushing teeth-a user study with the tebra system. ACM Trans Accessible Comput (TACCESS) 5:1–34. https://doi.org/10.1145/2579700

# Research on Security Awareness to Protect Data Through Ontology and Cloud Computing

**G. M. Kiran and N. Nalini**

## 1 Introduction

Cloud Computing is an advanced technology used for storing the data and its fast growth demands on numerous applications and multiple servers in the networking field [10]. The advantage of this technology is the data can be accessed and secured efficiently from any location at any time [11]. Ontology is another set of paradigm that can be used for representation of knowledge in the form of formal and structured [12]. Different data from various sources are stored in the network, therefore there is a need for security in the ontology which is very much essential [13]. For data security, basic user identification is the basic means that prevent unauthorized access to the cloud server [14]. Cloud Computing comprises three components namely the device owner, the user and the server client. The owner of the content resource will transfer the encrypted information to the server, where the required user removes the information [15]. During the retrieval process, the user is verified with his / her authentication policies to ensure that the information is in the right hands. Recently numerous strategies have been proposed for security in Clouds and Ontology, such as Fine-grained data access control, role based access control and so on, however those methods proposed are quite older and not efficient for storing the larger amount of data for longer period of time in the Ontology. Hence there is a need of an efficient security model to overcome the limitations of existing methodologies in the field of Cloud Computing.

G. M. Kiran (✉)
Shridevi Institute of Engineering and Technology, Tumakuru, Karnataka 572106, India

N. Nalini
NITTE Meenakshi Institute of Technology, Yelahanka, Bengaluru, Karnataka 560064, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_4

## 2 Review of Related Works

In recent years, many data access control with different security methods are developed, some of the works reviewed here.

Kayes et al. [15] developed a model in the medical field whereby accessing the data in the server side was a major concern since the unauthorized users can also access the data, To overcome the current issues, the ontology method is proposed. The security and privacy requirement can be achievable with the utilization of the ontology process.

Yang et al. [4] developed a model whereby multi server systems are difficult to deploy and operate in practice since either a user has to communicate simultaneously with the multiple server or the protocols are quite expensive. In proposed system only a front-end service can be used to communicate with the users while a control server will stay behind the scene, multi server password system must be proposed to overcome the problem of single point of vulnerability inherent in the single server architecture.

Leao et al. [6] distribution of key shares is one of the novel contributions of this paper, by adopting Secret Sharing Scheme in different administrative domains, but retrieving the information stored in the multiple systems was a formidable task.

## 3 Gap Analysis

| Reference No. | Proposed system | Future work (or) limitations |
|---|---|---|
| [1] | The focus of this paper is on grid programming to develop grid-enabled services using the Globus Toolkit. Assessment data suggest that undergraduate computer science students can specialize in grid programming with preparation equipment and manual exercises | The Virtual Data Toolkit should be made available for installation of the globus toolkit environment |
| [2] | The reliability of the Grid Computing Platform is guaranteed by a trusted virtual machine, and only the computing platform that satisfies the Grid's secure system can access the Virtual Organization through the technology of Trusted Network Connection (TNC). Meanwhile, the insecure platform is torn | The spread of viruses and malicious software on the grid should be effectively prevented and other grid security issues addressed |

(continued)

(continued)

| Reference No. | Proposed system | Future work (or) limitations |
|---|---|---|
| [3] | This paper introduces confidentiality technology, which uses encryption, and relates to the construction of data and their organizations, providing a natural way to promote authorization and as a framework for many use cases | Grid is compatible with middleware and provides security for server-side data, to provide a model that can be deployed in different environments |
| [4] | A password-based user authentication system is created to protect the data stored on a single server | A multi-server password plan should be proposed to prevent a single-site vulnerability from being built on a single server |
| [5] | GOS aims to support the services of centralized data banks and data sources that can be shared between multiple computers and end users in the grid | Recovering the information stored in many programs was a huge undertaking |
| [6] | This paper proposes a method for automatically studying the well-established ontologies described in UML by using word classification | Establish an automated way of generating well-established ontologies by using language concepts such as semantic types |
| [7] | Proposed a method of ontology development using a map-based visualization technique | It is the intention of the future to add a collaborative phase of ontology content prior to implementation this will enhance the ontology quality confidence and its reuse |
| [8] | Explains the function of the visual and theoretical frame to allow better user control over data acquisition data | Ways to provide a more up-to-date presentation of data should be improved |
| [9] | New document semantic ranking process for the semantic ranking that proposes a new weight of the query in the document based on computer science ontology weight | Use the proposed method in semantic search applications using Computer science ontology and show the results using the information view method |

## 4 Problem Identification

The current security infrastructure on file servers is simple. Therefore, there is a need for an efficient approach to file data security infrastructure. Security and privacy were the most worrying issues in the current system. These servers are built on more than one single server architecture. Data security is an important part of these servers, but these file servers result in one vulnerability increasingly compromised. Therefore, the focus of the problematic statement of work is not just to hide information but also to protect it. The focus of the work is to design a model for effective conservation and management of data captured in distributed environments.

## 5 Contribution and Organization of the Research

The proposed system includes the following contributions:

1. The proposed process was developed to achieve security especially in the field of Cloud computing using Ontology.
2. The proposed process is streamlined to achieve security improvements and guaranteed control of data access for authorized users.
3. The proposed procedure is developed to properly store the most long-term data in the field of medicine.

## 6 Concepts of Ontology

Generally, Ontology is defined as "a branch of metaphysics relating for establishing the relationships among each other in an organization".

The basic structure of Ontology and its access control is illustrated in Figs. 1 and 2 as shown.

The Ontology has 3 basic models as shown below:

1. The Subject actually plays as a resource owner, who is responsible for creating and defining the rules among the users.
2. Policies are nothing but the rules used for accessing collection of information which are actually generated by the resource owner.
3. The Object is a collection of information such as files which are used for transmissions.

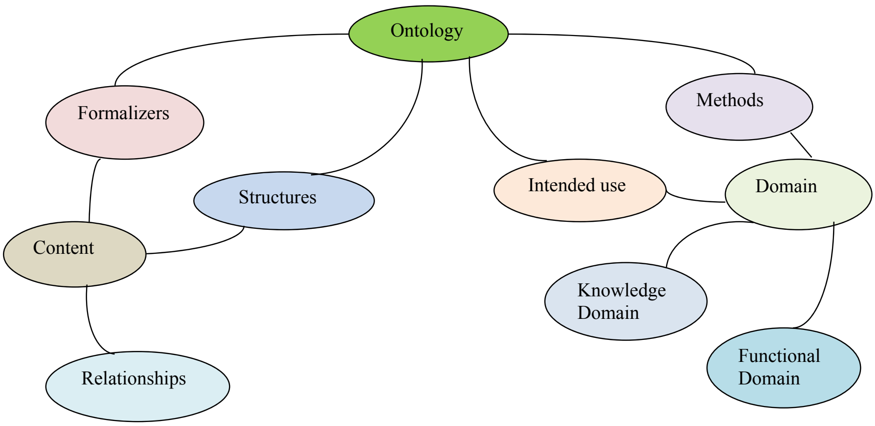



**Fig. 1** Basic structure of ontology

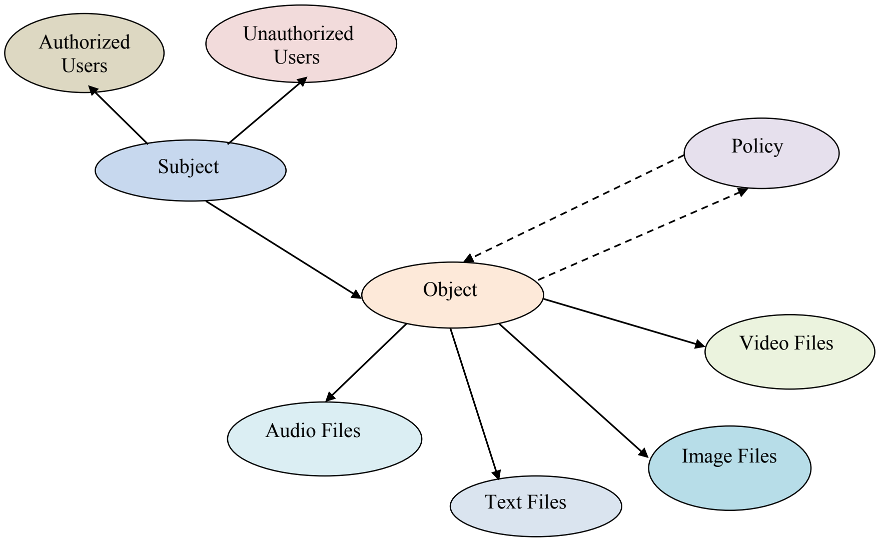



**Fig. 2** Structure of access control in ontology

**Table 1** Analysis of the proposed method parameters

| S. No. | Parameters | Data in MB | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 10 | 20 | 30 | 40 | 50 |
| 1 | Delay | 0.3217 | 0.5879 | 0.7943 | 1.0935 | 1.3498 | 1.5497 |
| 2 | Encryption time | 0.2918 | 0.5379 | 0.7353 | 1.1345 | 1.3979 | 1.5998 |
| 3 | Decryption time | 0.0248 | 0.0298 | 0.0312 | 0.0289 | 0.0439 | 0.0937 |
| 4 | Information retrieval time | 0.114 | 5.9281 | 11.625 | 20.139 | 36.148 | 55.458 |
| 5 | Processing time of ontology | 2.938 | 2.9103 | 2.9259 | 3.147 | 3.343 | 3.947 |

## 7 Performance Analysis

See Table 1.

## 8 Conclusion

User privacy and data security are a major concern for network development. New innovations are being developed to improve the security through the dissemination of new security policies. As noted in previous research, the build of a single server is what creates these security gaps in the company network, thus being vulnerable to all server data. This research work has the potential to protect data by storing data on multiple back-end servers through the ontological infrastructure used to store and retrieve data in an efficient and effective way. As seen in previous researches, the single server architecture is the one who are creating these security loopholes in the corporate network, thereby posing danger to the entire server data.

This Audit function has the ability to secure server data by storing data on small n servers running in the back-end using the ontological infrastructure used to store and retrieve data in an efficient and effective way. In this system only the primary risk server deals directly with users while the underlying servers stay behind the scene, and therefore can work directly to strengthen the build of a single server. The simulation was started by illustrating the method in Mat Lab and performed under standard conditions. In the end, the result proves a viable strategy for achieving the best performance of medical ontology in a sequence.

## References

1. Mache J, Apon A (2007) Teaching grid computing topics, exercises and experiences. IEEE Trans Educ 50:3–9
2. Wang Z, Feng Q, Xu R (2008) Research on trusted access technology of grid resource based on the virtual machine. In: IEEE international conference for young computer scientists, pp 1384–1388
3. Blanquer I, Hernandez V, Segrelles D, Torres E (2009) Enhancing privacy and authorization control scalability in the grid through ontologies. IEEE Trans Inform Technol Biomed 13:16–24
4. Yang Y et al (2006) A practical password-based two-server authentication and key exchange system. IEEE Trans Dependable Sec Comput 3
5. Zhigangwang F, Helian N et al (2007) Grid-oriented storage: a single-image, cross-domain, high-bandwidth architecture. IEEE Trans Comput 56:474–487
6. Leao F, Revoredo K, Baiao F (2013) Learning well-founded ontologies through word sense disambiguation. In: IEEE conference on intelligent systems, pp 195–200
7. Iqbal R, Murad MAA et al (2013) An ontology engineering approach with a focus on human centered design. In: IEEE conference on intelligent system design and applications, pp 86–90
8. Yager RR, Petry FE (2006) Multicriteria approach to data summarization using concept ontologies. IEEE Trans Fuzzy Syst 14:767–780
9. Boonyoung T, Mingkhwan A (2014) Semantic ranking based on computer science ontology weight. IEEE, pp 86–91
10. Auxilia M, Raja K (2016) Ontology centric access control mechanism for enabling data protection in cloud. Indian J Sci Technol 9(23):1–7
11. Bokefode JD, Bhise AS, Satarkar PA, Modani DG (2016) Developing a secure cloud storage system for storing IoT data by applying role-based encryption. Procedia Comput Sci 89:43–50
12. Choi C, Choi J, Kim P (2014) Ontology-based access control model for security policy reasoning in cloud computing. J Supercomput 67(3):711–722
13. Choi C, Choi J, Ko B, Oh K, Kim P (2012) A design of onto-ACM (ontology-based access control model) in cloud computing environments. J Internet Serv Inf Secur 2(3 & 4):54–64
14. Imran-Daud M, Sánchez D, Viejo A (2016) Ontology-based access control management: two use cases. In: Proceedings of the 8th international conference on agents and artificial intelligence, pp 244–249
15. Kayes AS, Han J, Colman A (2015) An ontological framework for situation-aware access control of software services. Inf Syst 53:253–277

# Driver Activity Monitoring Using MobileNets

**Deval Srivastava, Priyank Shah, and Saim Shaikh**

## 1 Introduction

In our current world as driving technology continues to grow the driving effort required decreases. Hence, drivers become more and more careless resulting in loss of life in many circumstances. The proposed system aims to solve this problem by developing a system to monitor the driver's activity and warn them whenever necessary. The method involves deploying a neural network trained on various categories such as talking or texting on the phone, talking to co-passengers, operating the radio, and drinking water. In our paper, we have extensively tested the performance of different neural networks [1] such as Resnet-50 [2], Inception [3], and MobileNets [4]. Throughout the development, our focus has been to make a system that replicates the driver's real-life conditions. Hence, our network will receive the images from an IR camera allowing our system to perform during nighttime. Our system can be easily fitted to any existing vehicle very easily and will be intuitive to use. Our system has been developed such that it can work even in regions having extremely poor internet connectivity. The system will also be equipped with sensors to detect rash driving and will consist of security features such as fencing, fingerprint authentication to prevent thieving of the vehicle.

From the previous work and research done in this domain, it can be concluded that the most popular computer vision methods include detecting driver inattention using head pose, eye gaze estimation or simply checking eye closure rate as well as measures such as EEG, electrocardiogram, etc. We will discuss these methods and other techniques that have been used in the next section.

The paper has been organized in the following manner next we will be looking at some of the most comprehensive research that has been done in this field and is relevant to this application, Further that the algorithm and its peculiarities will be

D. Srivastava (✉) · P. Shah · S. Shaikh
Fr. Conceicao Rodrigues College of Engineering, Mumbai, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_5

explained, after that we can take a look at the complete flow of the system and our design methodology. Post that the dataset, training, and parameters will be discussed. Towards the end, we can look at the results received by us and the conclusion.

## 2 Literature Survey

According to the survey conducted by us, it can be concluded that the most popular methods to solve this problem involve either driver biological measures, driver physical measures, driving performance measures, or some kind of a hybrid measure [5].

Driver biological measures include biological signals like EEG, electrocardiogram (ECG), electro-oculography (EOG). These signals are collected through electrodes in contact with the skin and then analyzed for fatigue and drowsiness. Physical measures involve eye closure detection and blink frequency, face position, driver gaze to detect inattention.

Driver performance measures involve various measures such as steering angle and other driving criteria. Most research that has been done in the related field has been focused on detecting driver inattention using eye gaze tracking and head pose estimation. These methods rely only on the head and eye movement to detect inattention whereas in real life a driver can be distracted doing various tasks that cannot be detected by head movement alone. It has been observed that current driver monitoring systems employ statistical machine learning methods to detect driver distractions and work on a limited dataset. Research done by Martin et al. [6] involves classifying drivers' gaze into various regions using a machine vision algorithm that utilizes face detection and facial landmark extraction to estimate the gaze of the driver further they are using that to explore driver's gaze dynamic patterns. Further, the authors have condensed gaze dynamics into glance frequencies and duration.

Some research has been done on applying deep learning technologies to solve this problem but such systems cannot be cost-effectively deployed in a vehicle nor do they work in nighttime conditions. These methods have used architectures like the vgg-16 and have been trained on datasets like the StateFarm dataset [7].

## 3 Algorithm

For the development of our system, we made use of the MobileNet Algorithm [8]. MobileNet [8] is a neural network that was developed by Google to perform on low powered devices lacking graphical GPUs that are known to accelerate neural network performance. MobileNets are small, low-latency, low-power models parameterized to meet the resource constraints of a variety of use cases, one of those use cases is that it can also be deployed on a Raspberry pi which we intend to do.

A standard convolution, filters and combines inputs into a new set of outputs in one step, but in the case of MobileNets it first uses depthwise convolution [8] that applies a single filter to each input channel. The pointwise convolution then applies a $1 \times 1$ convolution to combine the outputs of the depthwise convolution. The depth wise separable convolution splits this into two layers, a separate layer for filtering and a separate layer for combining. This factorization has the effect of drastically reducing computation and model size, this modification allows MobileNet to be faster than its other counterparts.

For our application we have employed a MobileNet v2 [8], it's the second iteration of MobileNets and now along with the depthwise separable blocks it also uses bottleneck residual layers and also adds a $1 \times 1$ expansion block whose purpose is to expand the number of channels in the data before it goes to the next block. In the proposed system we have used MobileNet v2 as it's much better than its older version, In our mobilenet model we have 17 bottleneck layers which are followed and preceded by convolutional 2d layers and a total of 3.4 M parameters which to put in perspective are far less than 138 M parameters of VGG-16 and 4.2 M parameters of the competing mobilenetV1. The model was on a self-made dataset of drivers performing distracting, reckless activities.

# 4 Proposed Method

## *4.1 Implementation*

For lucid understanding of the implementation we have developed, the paper presents it in points listed below.

1. Figure 1 describes the entire workflow of the project right from the hardware setup to the user interface. The Pi camera is mounted on an appropriate position in the dashboard of the vehicle. It is then connected to the camera port on the Raspberry Pi. The SM808 GSM + GPS module is connected to the Raspberry Pi via USB TO RS232 serial port. The GPS antenna is connected to the module and placed outside the vehicle with a clear view of the sky. The Raspberry Pi is then connected to a portable power supply via the micro USB port.
2. The Pi camera records footage of the driver and sends the frames to the Preprocessing Unit. The footage is recorded at a resolution of 640*480 at 24 frames per second. The preprocessing unit then performs basic image processing, noise reduction on every third frame and resizes them to 224*224*3 then it is forwarded to the neural network which predicts the class of that image and depending on that result we declare the driver as distracted or not. If the driver is distracted the buzzer is rung to alert the driver.
3. As similar performance is targeted for both night and day time footage. In low light, the Pi Camera is aided by 2 IR bulbs which help provide clear frames even in pitch black conditions. The frames received from the Pi Camera are

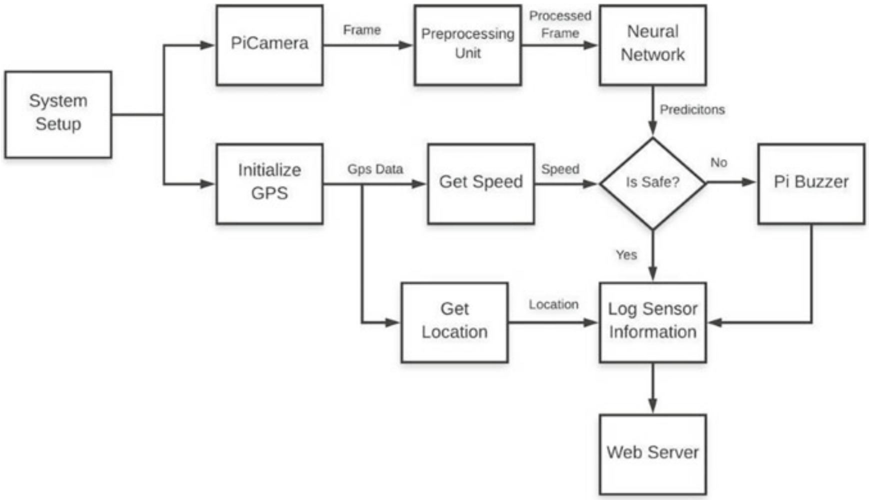



**Fig. 1** Implementation flowchart

processed. Alongside this, we also calculate the speed of the vehicle and find out the location of the vehicle the speed is checked to be within the prescribed speed limit and the location within bounds, if either of these details is found to be dissatisfactory then they are reported to the administration with an alert. Once both operations are complete data is pushed to the administrative web server.

## *4.2 Dataset*

When we were looking for a dataset that would be able to suffice our needs for the classification tasks we came across many publicly available datasets one of them being the NTHU Driver Drowsiness Detection Dataset [9]; this dataset consisted of different subjects performing a variety of tasks both during driving many of which can be considered as distracted in our system however one shortcoming that we landed upon while we worked in this dataset was the fact that the dataset was not recorded in an actual car and that factor would affect the performance of the system in real-world cars and secondly this dataset lacked any data points for certain classes we held valuable and were crucial to our research these were the drinking class and use of a smartphone. Overuse of smartphones and drinking while driving are both issues relevant to our current day society. The second dataset that we found was the Kaggle Statefarm Dataset [10]; This dataset included photos of various subjects performing activities in a vehicle where some of them could be considered as reckless and distracting however one of the flaws of this dataset was the lack of any infrared imagery for nighttime and just very poor positioning of the camera

**Table 1** Database statistics

| S. No. | Class | Count (day + night) |
|---|---|---|
| 1 | Safe driving | 5000 |
| 2 | Talking on phone | 5000 |
| 3 | Texting on phone | 5000 |
| 4 | Drinking | 5000 |
| 5 | Sleeping | 5000 |
| 6 | Yawning | 5000 |

system which would be virtually impossible in a real-world car or truck. While evaluating publicly available datasets we deliberated over each of our options and even considered combining the datasets but none of these approaches proved to be satisfactory as our neural networks trained on these datasets were not performing as expected in real-world conditions [11]. After this realization, we started working on creating a real-world dataset curated to our task. We recorded 6 different drivers performing various distracted activities across multiple cars. We recorded drivers performing activities such as talking on the phone, texting while driving, drinking, sleeping, Yawning. All of these activities were performed in simulated environments where the drivers are not driving. The Dataset was recorded using a pi camera as it's the camera that will be feeding images to the neural network. We recorded images for the night dataset by using an IR camera and IR lights. The database statistics are presented in Table 1 (Figs. 2 and 3).

In the above pictures, we can see the dataset samples from the night and the day. The images are in order of Sleeping, Talking on the phone, Drinking, Texting, Yawning.

**Fig. 2** Dataset samples that were taken during the day

**Fig. 3** Dataset samples that were taken during the night

## 4.3 Training and Configuration

This section will discuss how we trained our model on the above dataset. Now coming to train our model we decided to use a transfer learning approach this method says that first, the model has trained on a huge dataset such as the imagenet [12] dataset for image classification tasks, it contains 1.4 M images and 1000 classes allow the model to learn features of images which becomes our base knowledge. The model which has learned these image features now can be carried over and used for other computer vision tasks, this is done precisely by capturing the weights of all the layers but the bottom few fully connected layers, the top layer weights are said to be 'frozen' ie when the model has trained on the new dataset or in this case our dataset the weights will not change and the backpropagation and weight changes will be limited to the bottom layers. Transfer learning allows neural networks to train better and quicker while promoting reusability and modularization.

In our case, we employed a mobilenet V2 model that had been trained on the imagenet dataset further to utilize this model in our application we choose the freeze the bottleneck layers of the model as they are more generalizable and add a layer to average the weights and then a fully connected layer with the softmax activation that will give us the logits.

We have used transfer learning to train our neural network as we have used the model of MobileNet v2 for feature extraction and only trained the last few layers to get the best results and quicker training times. We used 60%, 20%, 20% split for training, validation, and testing respectively for our model. To further simulate real-world conditions we added data augmentation to our model. This allows our model to perform better in difficult scenarios like low lighting, improper camera alignment, etc. After experimenting and testing with a lot of different kinds of augmentations we found the following augmentations gave us the best results were random zoom that generates extra images that are zoomed in randomly up to 20%. In the same way, we added random crops up to 20% and also random brightness for varied conditions. Augmentations effectively increase the size of our dataset and also makes sure the model works better in unknown conditions.

We trained the MobileNet v2 model on our custom dataset with a batch size of 32 with Adam optimization [13] which is known to perform well on similar tasks with a parameter of 0.01, we trained the model on a computer with an Nvidia GTX1070TI with Cuda acceleration. Our model was trained for roughly 6500 steps with a batch size of 32 which took 4 h to train and we stopped when we had a validation accuracy of 95.6%.

## 5 Results and Discussion

In this section, we will discuss the results we have received after implementing our application to completion. Let's first understand how results are organized, In Table 1 we are looking at all the models we have trained on our datasets to measure up performance and other factors, we have recorded the accuracy of the model on our test split of the dataset, the average CPU usage recorded over 6 h of the utilization of our system and the average frames per second that we were able to process on our system. Further, we also trained our model on various other architectures such as Inception v2, ResNet-50, and VGG-16 [14] model on the same dataset to compare the performance we receive on the Raspberry pi and give a comprehensive result. In a cursory analysis of the table, we can observe that mobilenetV2 offers the highest frames per second and the lowest CPU utilization to go along with it, looking at fps and accuracy metrics of other models we can infer that there is accuracy vs performance tradeoff present here. Higher parameter count and more computationally heavy models such as inception offer higher accuracy to go with it. Coming to our use case there were potential concerns and requirements we had that helped us navigate around the tradeoff. We are deploying the model on a raspberry pi, which is a rather small and weak computer, and along with processing driver images through a neural network, we were also doing auxiliary processing which includes velocity and GPS information calculation, sending regular updates to an administration webserver. With this much processing to go around it becomes crucial to manage CPU resources which are already scarce in raspberry pi. Next concern was regarding the speed of processing frames, in situations where the driver may fall asleep on the wheel we need to be able to act fast and quickly process the data so the driver can be alerted and an accident can be averted, with this in mind it becomes paramount to have high frame per second as to decrease the time spent on other frames before reaching the one deciding frame. Now with consideration to these concerns, we can see that MobileNet v2 is the most ideal neural net model for our use case. We believe that MobileNet v2 was the best model for our application as we want to have the system to be almost real-time which will enable it to prevent accidents. As MobileNet v2 requires the least amount of CPU usage some processing ability of the limited compute on a Raspberry pi CPU can also be used

**Table 2** Comparison of different models

| S. No. | Model name | CPU usage (%) | Fps | Execution time (min) | Accuracy (%) |
|---|---|---|---|---|---|
| 1 | Inception v2 | 90 | 1 | 8 | 98 |
| 2 | VGG-16 | 93 | 1 | 8 | 96 |
| 3 | Resnet-50 | 91 | 2 | 4 | 97 |
| **4** | **MobileNet v2** | **66** | **6** | **1.34** | **93** |

for other tasks such as calculating speed and sending data to a web server. We further tested our system in real-world conditions and found satisfactory results (Table 2).

In the execution period column, we have provided the time it took to run a 60 s video which is encoded with 24 fps and we process every one in three frames. In the second section, we can look at and understand the results we have received in our real-world testing. Firstly, the metric we have developed to compare and analyze results from ours and other competitive solutions in an equitable manner is explained; we recorded over 100 instances of distracted driving doing various activities in both day and nighttime conditions. Next, we ran existing solutions developed for this problem to compare results with our solution. We have chosen solutions utilizing deep learning, eye closure estimation and gaze estimation from head pose for comparison. The deep learning method utilizes a VGG based architecture and has been trained on the NTHU dataset. Since statistical methods involving eye/head estimation cannot generally provide you with the exact distracted activity but rather binary information about whether the driver is focused or not nevertheless we will be using that only for our comparisons. To have fair results we will be considering classes from the existing deep learning solution which are also present in our solution. We believe the sleeping/sleepiness category is an essential element that warrants it to have a separate category.

In Table 3 we can look at the results of each method on detecting other distracting activities and sleeping/sleepiness, we are comparing percentage-based results to assess the number of times each method detected the activity correctly. In a cursory glance, we can infer from this table that in sleeping/sleepiness category are method edges out ahead and moves ahead in the other category with ease, looking at the

**Table 3** Comparison with existing solutions

| S. No. | Method name | Sleeping/sleepiness (%) | Other distracting activities (%) |
|---|---|---|---|
| 1 | Deep learning-based method | 90 | <u>85</u> |
| 2 | Eye closure estimation method | 88 | 68 |
| 3 | Gaze estimation from head pose method | <u>92</u> | 72 |
| **4** | **Our method** | <u>**95**</u> | **93** |

underlined entries we can understand from that a deep learning solution performs better on detecting activities but falls short of head pose estimation methods on detecting drowsiness which does well on that task, however, our deep learning solution trained on a custom dataset successfully manages to outperform existing deep learning solutions on activity detection and edges out ahead in drowsiness detection compared to head pose estimation.

## 6 Conclusion and Future Scope

On successful implementation the system will provide a robust and efficient method to monitor driver activities and thus prevent accidents that occur due to distracted driving, overuse of mobile phones while driving, texting on the phone, drowsiness, sleeping, etc. When such a system is in place it will enforce the drivers to be more careful and drive responsibly which will prevent loss of lives and will promote a safer driving experience for other drivers on the road. Further we can conclude from the results that implementation of mobilenet for this task allows our system to run more efficiently and much quicker compared to contemporary neural network architectures. The decision to go ahead with a custom dataset over open source readily available datasets allows us to add curated classes and more variability and real world features that proved beneficial.

Once applied over a large number of vehicles the system can also be used to create a network of vehicles to share important information. In the future, this network will be able to collect huge amounts of data and this data can be used to plan routes better. Moreover, since we have deployed a hardware platform more and more features can be added in due time. Features such as facial recognition for authentication and various kinds of analysis can be done using the data of our platform. The algorithm in the proposed system relies on a neural network to detect driver's activity which performs well but an object detection approach can be used to detect specific distracting objects which will theoretically perform even better than standard neural net approaches.

## References


1. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient based learning applied to document recognition. Proc IEEE 86(11)
2. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), Las Vegas, NV, pp 770–778. https://doi.org/10.1109/CVPR.2016.90
3. Szegedy C et al (2015) Going deeper with convolutions. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR), Boston, MA, pp 1–9
4. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017). MobileNets: efficient convolutional neural net

5. Ahir A, Gohokar V (2019) Driver inattention monitoring system: a review. In: 2019 international conference on innovative trends and advances in engineering and technology (ICITAET), Shegaon, India, pp 188–194. https://doi.org/10.1109/ICITAET47105.2019.9170249
6. Martin S, Vora S, Yuen K, Trivedi MM (2018) Dynamics of Driver's gaze: explorations in behavior modeling and maneuver prediction. IEEE Trans Intell Veh 3(2):141–150. https://doi.org/10.1109/TIV.2018.2804160
7. Vicente F, Huang Z, Xiong X, Torre F, Zhang W, Levi D (2015) Driver gaze tracking and eyes off the road detection system. IEEE Trans Intell Transp Syst 16(4)
8. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L (2018) MobileNetV2: inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, Salt Lake City, UT, pp 4510–4520
9. Weng C-H, Lai Y-H, Lai S-H (2016) Driver drowsiness detection via a hierarchical temporal deep belief network. In: Asian conference on computer vision workshop on driver drowsiness detection from video, Taipei, Taiwan, November 2016
10. Kaggle statefarm dataset. https://www.kaggle.com/c/state-farm-distracted-driver-detection
11. Valeriano LC, Napoletano P, Schettini R (2018) Recognition of driver distractions using deep learning. In: 2018 IEEE 8th international conference on consumer electronics—Berlin (ICCE-Berlin), Berlin, pp 1–6. https://doi.org/10.1109/ICCE-Berlin.2018.8576183
12. Deng J, Dong W, Socher R, Li L, Li K, Fei-Fei L (2009) ImageNet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, Miami, FL, pp 248–255. https://doi.org/10.1109/CVPR.2009.5206848
13. Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In: ICLR
14. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: International conference on learning representations

# Prediction of Crop Production Using Analysis Algorithms

**Arun Pratap Tomar and N. Nalini**

## 1 Introduction

India is an agrarian country yet GDP from agriculture in India is decreasing year by year. Farmers are not getting information about the crop-growing season in the early stage, especially in big agricultural areas. Sometimes getting the exact production after harvesting is also important.

The most challenging task for the farmers is to get the information about the crop growth in the respected area in the upcoming season and the forthcoming weather knowledge in that respected area which ultimately leads to the crop production statistics.

According to the traditional methods, farmers check the past production, market price in the last years and go for growing the crop which is going to provide them better production but sometimes end up being in a loss. This happens because they don't consider the other factors such as upcoming weather, nutrients in the soil, temperature, humidity, expected rainfall, etc. This traditional approach is harmful for large scale crop production and can not provide real-time and forecasting estimation of crops as global warming and industrialization are affecting the rainfall which plays a vital role in crop production.

In the traditional approach, there are informal seed systems, often exchanged for free through social networks that don't provide high yield and most importantly there is low-tech postharvest processing which affects the next seasonal crop which is to be grown.

The crop yield mainly depends on changing weather conditions. Prediction of crop yield is a big problem which farmers face. All farmers get worried about the

A. P. Tomar · N. Nalini (✉)
Computer Science and Engineering, Nitte Meenakshi Institute of Technology, NMIT Campus, Govindapura, Yelahanka, Bangalore 560064, India
e-mail: nalini.n@nmit.ac.in

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_6

crop they are going to expect after cultivation. In the ancient times, prediction of crop yield was done by keeping farmers' familiarity with their crop. Now, the agricultural data is massive and hard to predict the crop yield. It will be very valuable for us when we will be able to turn the data into information [1].

In this paper, weather forecasting is shown as in which area of the country high percentage of rainfall is there, how much approximate expected rainfall will be there, which crop is better to be grown in a specific area according to last 20 years results, what is the cause of decrement in the crop production of a specific crop. Different algorithms are being used to show the past production area wise and how rainfall is helping or affecting crop production in different parts of the country.

## 2 Research Methods

It includes dataset being used, data cleaning and aggregation, data manipulation and normalization and building prediction models using different simple and efficient tools.

**Dataset Used**

Two datasets are being used for this implementation.

1. Crop dataset
2. Rainfall in India 1901–2015

Datasets were gathered from the publicly available resources of the Indian government website mygov.in. The first dataset consists of year, production, area, state name, and second dataset consists of subdivisions of states and rainfall annually, monthly and quarterly.

- Rainfall (mm): Annual rainfall for all crops in India was considered.
- Area (Million Hectares): The total cultivated area for all seasons such as Rabi, Kharif, and the whole year is considered for study for all types of crops.
- Production (Million Tonnes): Production for the different cultivated crops of different states and areas for all seasons.

**Dataset Preprocessing**

Dataset which is used for study was prepared after merging the above two datasets together. It consists of different attributes such as year in which the specified crop is grown, Season, type of crop, specific area, production of crop, and annual_rainfall. Two datasets were merged together for better analysis and were stored.csv file.

**Data Cleaning and Manipulation**

Before applying different types of analysis patterns, data cleaning and normalization was done to make the dataset dimensionally reduced and suitable for operations.

In data cleaning, first, the null values were checked in the dataset and were filled by taking the mean of the entries of that respected column for production. After that correlation was checked among attributes using heatmap to extract the knowledge as which analysis pattern is to be used [2].

Now, after cleaning the data and getting a proper dataset different analysis patterns were used. Regression analysis techniques were used for this model. It is a predictive modeling technique that shows the correlation between target variables and independent variables. For this model, production was considered as a dependent variable and other attributes (crop_year, area, annual_rainfall, season) were considered as independent variables.

The regression and classification techniques used in this study were Linear Regression, Gradient Boosting Classifier, MLPClassifier, and Random Forest Classifier.

**Building model using Scikit-learn**

In this, the discussion of building the prediction model for Linear Regression, Gradient Boosting Classifier, MLPClassifier, and Random Forest Classifier was done. Scikit-learn is an open-source tool for predictive data analysis by using it as an interface with the python as it has in built libraries. It is licensed under a simplified BSD license and provides different analysis algorithms. Figure 1 shows the steps involved in building the prediction model.

## 3 Experiment Result [3]

**Merging of both datasets**

Both the datasets are merged together using the pandas library of python by using merge command as shown in Fig. 2.

**Correlation Using HeatMap**

For any analysis relation between the attributes should be known to select an appropriate algorithm. This Fig. 3 shows a heatmap which is used to know the correlation between the attributes.

**Average Rainfall**

Figure 4 shows a bar graph plot which is used to see annual rainfall for every state. Pictorial representation gives a better idea of data entries.

**Production State-Wise and Year-Wise**

Figures 5 and 6 are simple plotting of the production of the crops group by state and year. These plots had helped to visualize the data properly.

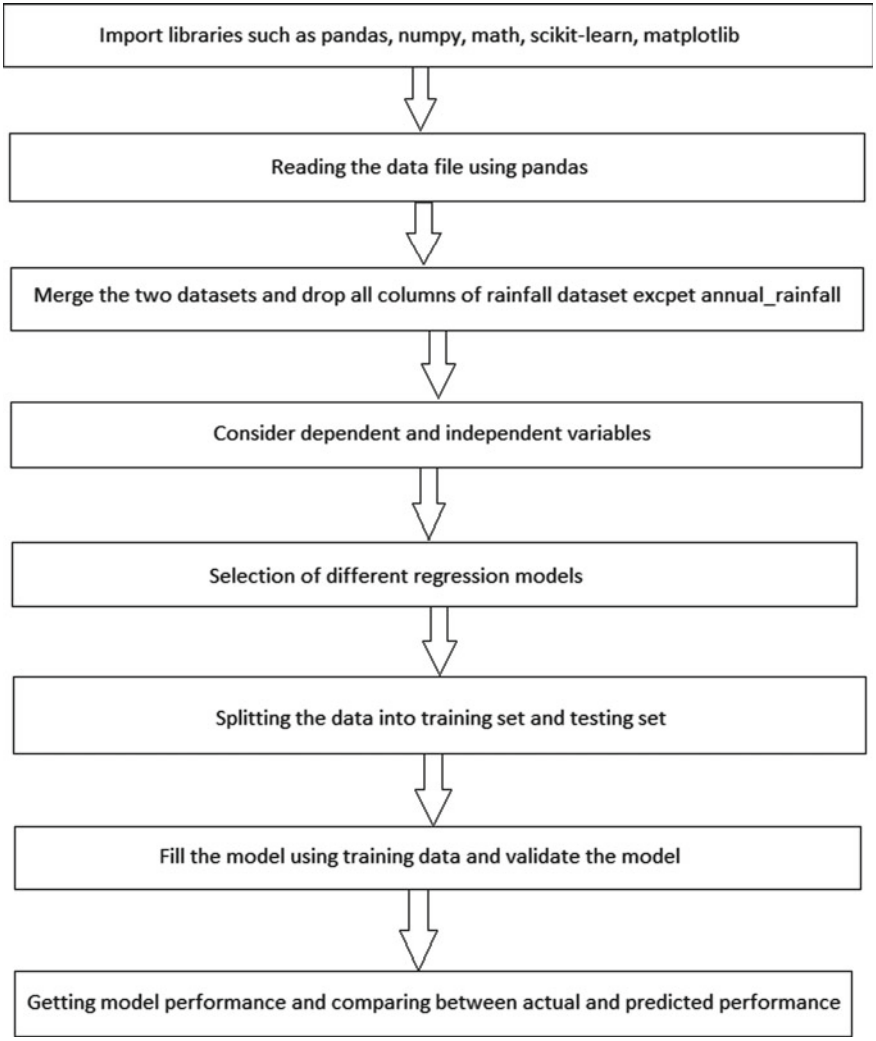



**Fig. 1** Steps involved in building the prediction model

**Scatter plot different types of Crops**

Figure 7 and Fig. 8 are mathematical diagrams which show the correlation between two variables as here for production and rainfall.

**Actual and Predicted Data**

Figures 9 and 10 shown below are plots which are displayed using linear regression to plot the residual between actual and predicted training and test data.

| | State | District_Name | YEAR | Season | Crop | Area | Production | Rainfall |
|---|---|---|---|---|---|---|---|---|
| 0 | Andaman and Nicobar Islands | NICOBARS | 2000 | Kharif | Arecanut | 1254.0 | 2000.0 | 2763.2 |
| 1 | Andaman and Nicobar Islands | NICOBARS | 2000 | Kharif | Other Kharif pulses | 2.0 | 1.0 | 2763.2 |
| 2 | Andaman and Nicobar Islands | NICOBARS | 2000 | Kharif | Rice | 102.0 | 321.0 | 2763.2 |
| 3 | Andaman and Nicobar Islands | NICOBARS | 2000 | Whole Year | Banana | 176.0 | 641.0 | 2763.2 |
| 4 | Andaman and Nicobar Islands | NICOBARS | 2000 | Whole Year | Cashewnut | 720.0 | 165.0 | 2763.2 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 67389 | Uttarakhand | UTTAR KASHI | 2014 | Whole Year | Onion | 26.0 | 172.0 | 1287.4 |
| 67390 | Uttarakhand | UTTAR KASHI | 2014 | Whole Year | Pulses total | 5346.0 | 4630.0 | 1287.4 |
| 67391 | Uttarakhand | UTTAR KASHI | 2014 | Whole Year | Rice | 9919.0 | 16499.0 | 1287.4 |
| 67392 | Uttarakhand | UTTAR KASHI | 2014 | Whole Year | Sunflower | 3.0 | 3.0 | 1287.4 |
| 67393 | Uttarakhand | UTTAR KASHI | 2014 | Whole Year | Total foodgrain | 35417.0 | 49689.0 | 1287.4 |

67394 rows × 8 columns

**Fig. 2** Merging two datasets together

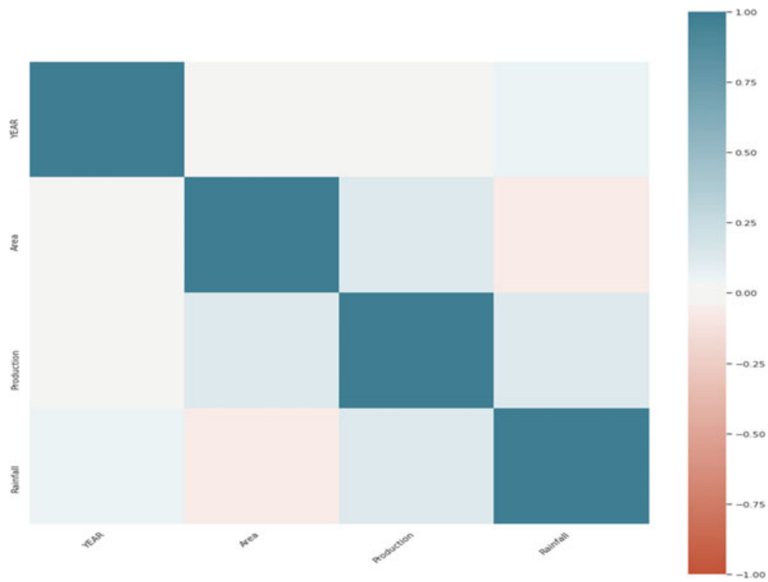



**Fig. 3** Heat map to show correlation between attributes

## Outputs of Different Analysis Techniques

(a) **Linear Regression**

This technique is a linear approach to model the relationship between the target variable and the remaining variables which help in achieving the target. When we use this technique only for one dependent variable it is called simple linear regression. When we have more than one target variable that time the linear

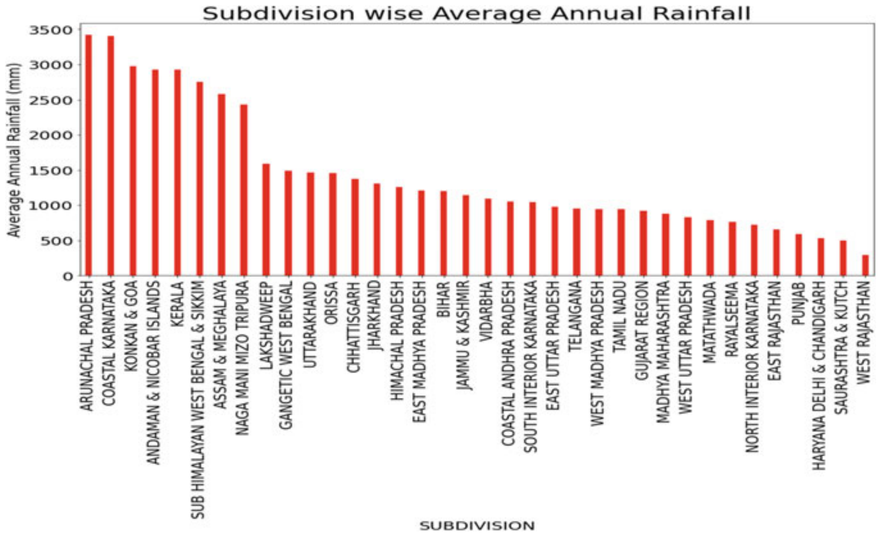



**Fig. 4** Graph showing annual rainfall state wise

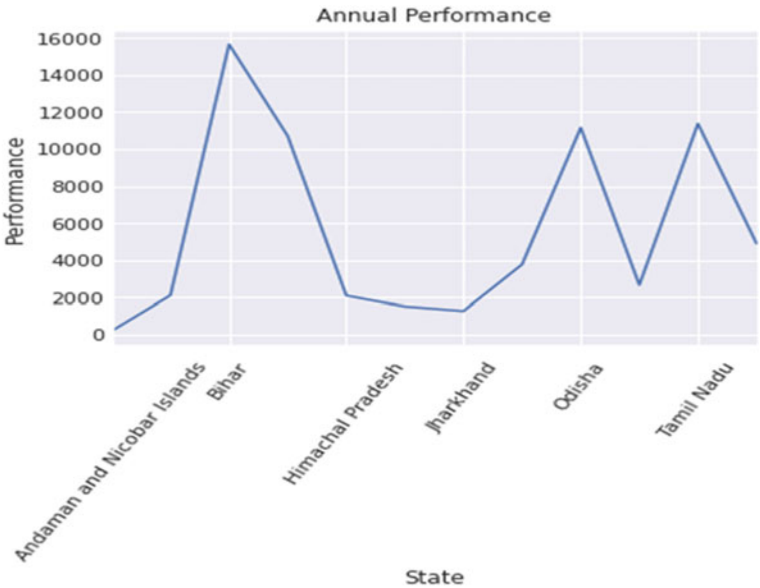



**Fig. 5** Performance of crop state wise

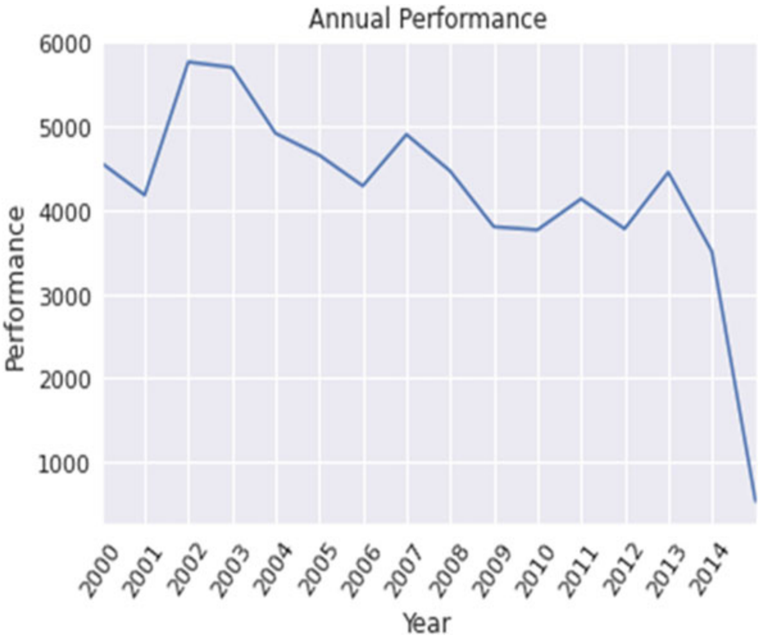
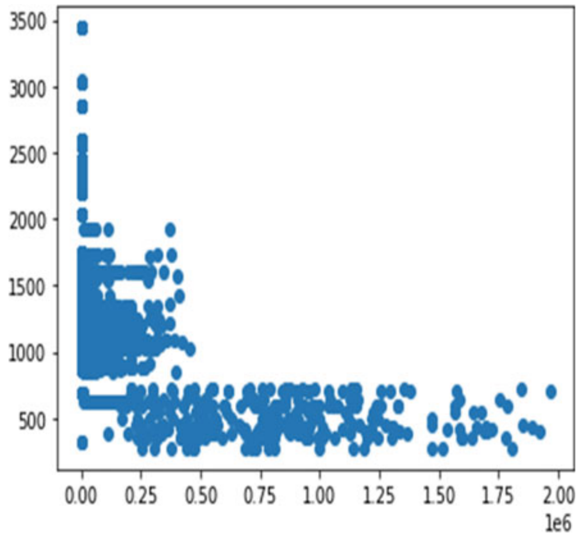



Fig. 6 Performance of crop year wise

Fig. 7 Scatter plot for wheat data



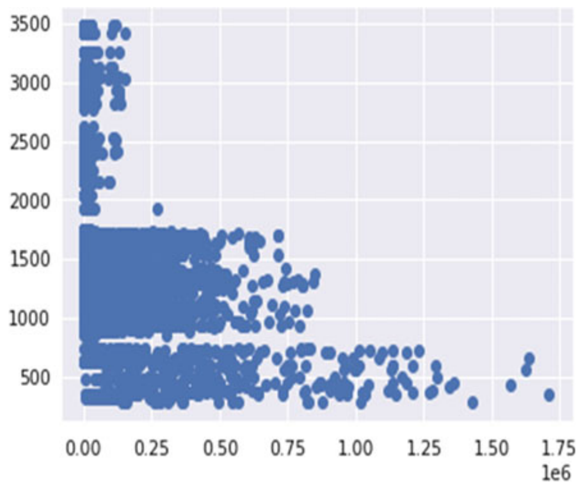



**Fig. 8** Scatter plot for rice data

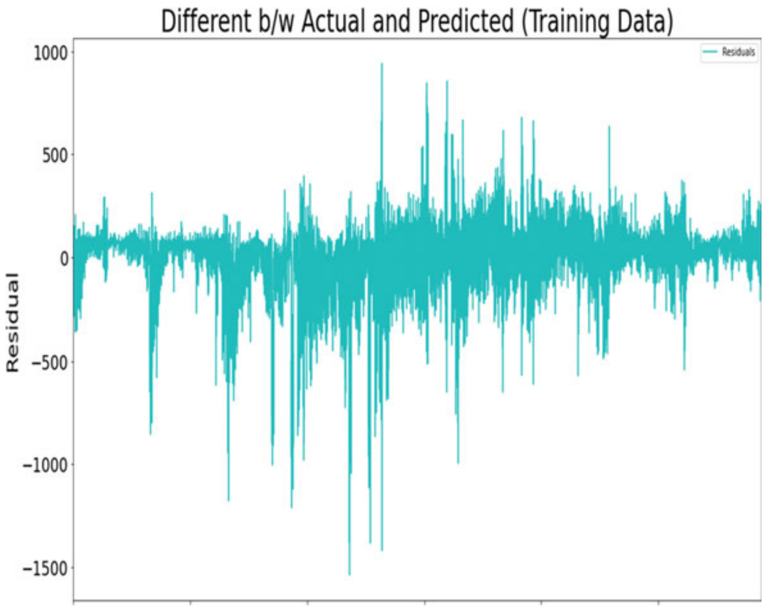



**Fig. 9** Residual of training data

technique is called multiple linear regression. It uses least square approach and conditional probability distribution for prediction [4].

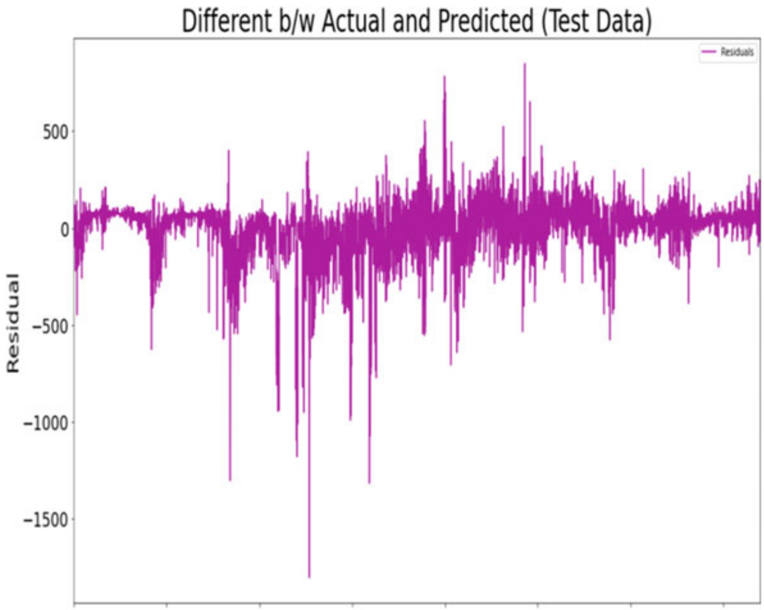



**Fig. 10** Residual of test data

```
[185] predicted = model.predict(test_x)
      model.score(test_x, test_y)

      0.9185399510349432

[186] from sklearn.metrics import roc_auc_score
      probabilities = model.predict_proba(test_x)

[187] roc_auc_score(test_y, probabilities[:, 1])

      0.9447971344009174

[187]

[188] from sklearn.metrics import precision_score

      train_predictions = model.predict(train_x)
      precision_score(train_y, train_predictions)

      0.8988744884038199

[189] from sklearn.metrics import recall_score

      recall_score(train_y, train_predictions)

      0.5853414769572459
```

(b) **Gradient boosting**

It is an ML technique for different types of analysis problems such as classification and regression. It predicts the model by forming different decision trees and combining them from the starting (It forms the first decision tree, then another and combines them to form next decision trees and so on to complete the prediction model with more accuracy). It builds the model by using stepwise technique and generalizes them [5].

```
[178] predicted = model.predict(test_x)
      model.score(test_x, test_y)

      0.9277394465464797

[179] from sklearn.metrics import roc_auc_score
      probabilities = model.predict_proba(test_x)

[180] roc_auc_score(test_y, probabilities[:, 1])

      0.9547685083277713

[180]

[181] from sklearn.metrics import precision_score

      train_predictions = model.predict(train_x)
      precision_score(train_y, train_predictions)

      0.844922592495408

[182] from sklearn.metrics import recall_score

      recall_score(train_y, train_predictions)

      0.7151582454192116
```

(c) **MLPClassifier**

A perceptron is a class of artificial neural networks (ANN). Single layer perceptron is an algorithm which learns only when patterns are linearly separable while multilayer perceptron (MLP) with two or more layers learns from all types of patterns and its processing power is very high. MLP contains at least three layers: Input, Output and Hidden Layer. Number of hidden layers depends on our model. All nodes in;ayers are non-activation functions. MLP uses supervised learning techniques for training the data and together input nodes with activation functions provides the target nodes (output layer nodes) [6].

```
[191] predicted = model.predict(test_x)
      model.score(test_x, test_y)

      0.8740262630758958

[192] from sklearn.metrics import roc_auc_score
      probabilities = model.predict_proba(test_x)

[193] roc_auc_score(test_y, probabilities[:, 1])

      0.9493502917255336

[193]

[194] from sklearn.metrics import precision_score

      train_predictions = model.predict(train_x)
      precision_score(train_y, train_predictions)

      0.5805435499230825

[195] from sklearn.metrics import recall_score

      recall_score(train_y, train_predictions)

      0.8800666296501943
```

(d) **Random Forest**

This technique is used for types of analysis algorithms classification and regression. It also forms the model by combining different decision trees like gradient boosting but aggregates the decision trees at the end. It decision the best decision tree and aggregates to predict the model hence reduce the overfitting and provides better accuracy when attributes are not correlated [7].

```
[ ] predicted = model.predict(test_x)
    model.score(test_x, test_y)

    0.9051858446472291

    from sklearn.metrics import roc_auc_score
    probabilities = model.predict_proba(test_x)

[ ] roc_auc_score(test_y, probabilities[:, 1])

    0.9274903392833863

[176]

[ ] from sklearn.metrics import precision_score

    train_predictions = model.predict(train_x)
    precision_score(train_y, train_predictions)

    0.9961486180335297

[ ] from sklearn.metrics import recall_score

    recall_score(train_y, train_predictions)

    0.9765685730149917
```

## 4 Conclusion and Future Work

Different analysis techniques were used for designing the model which can predict the crop production. Algorithms namely Linear Regression, Random Forest Regression, and Ridge Regression were used to build the predictive model. The experiment results showed that the Random Forest Classifier model is giving better accuracy than other analysis techniques such as Gradient Boosting Classifier, and MLPClassifier on the Potato, Wheat and other crops' dataset.

The results of these analysis algorithms show the precision and recall accuracy of the predictive model and it will show the percentage of chances the production will be more than 10,000 million tones in the respected state according to the rainfall in the future year.

Future work will be on the prediction of the statistics of the production for every respected year according to the state and predicted rainfall in that given year. Using different regression algorithms to analyze the dataset and apply different techniques to find the avg production of the respected crop and season state-wise as well as district-wise.

## References

1. Sujatha R, Isakki P (2016) In: International conference on computing technologies and intelligent data engineering (ICCTIDE'16). A study on crop yield forecasting using classification techniques
2. Shikha SJ, Saquib Z (2018) In: Fourth international conference on computing communication control and automation (ICCUBEA). Prediction of crop production in india using data mining techniques
3. Avouris N, Sgarbas K, Paliouras V, Koukias M (2017) In: IEEE global engineering education conference (EDUCON). Work in progress: an introduction to computing course using a python-based experimental approach
4. Zhang C, Wang F (2010) In: The 2nd international conference on industrial mechatronics and automation. Research on correlation analysis between test score and classroom attendance based on the linear regression model
5. Xu L, Liu J, Gu Y (2018) In: 10th International conference on modelling, identification, and control (ICMIC). A recommendation system based on extreme gradient boosting classifier
6. Singh J, Banerjee R (2019) In: 3rd International conference on computing methodologies and communication (ICCMC). A study on single and multi-layer perceptron neural network
7. Chippa M, Priyadarshini A, Mohanty R (2019) In: IEEE International conference on intelligent techniques in control, optimization and signal processing (INCOS). Application of machine learning techniques to classify web services
8. Gandhi N, Armstrong LJ, Petkar O, Tripathy AK (2016) In: 13th International joint conference on computer science and software engineering (JCSSE). Rice crop yield prediction using support vector machine

# A Deep Learning Approach for Speed Bump and Pothole Detection Using Sensor Data

**Bharani Ujjaini Kempaiah, Ruben John Mampilli, and K. S. Goutham**

## 1 Introduction

A majority of Indian roads are not monitored and maintained as per regulations of the concerned authorities. This poses a great obstacle for smooth and safe commuting. Municipalities like the Bruhat Bengaluru Mahanagara Palike spend crores of rupees to maintain and repair the roads of the city [1]. But these investments often don't bear fruit and common people spend a lot of money to make up for the damage caused by these ill conditioned roads. Bad roads result in severe accidents, injuries and loss of life. Keeping our road net- work in sound condition is a tough task due to a variety of factors such as harsh traffic densities [2], bad weather and extreme wear and tear brought about by extensive vehicle use. The two major road anomalies are Speed Bumps and Potholes [3, 4]. Methods to fill up the existing potholes are not very helpful and patched-up holes start wearing off within a couple of months. Unmarked speed bumps [5] cause significant inconvenience to commuters. If there could be a way by which the concerned authorities could be informed about aggravating road conditions before it has reached a state of total despair, it would play a major role in decreasing the number of accidents that take place every day. Moreover, early notification would also result in lowered cost for fixing a pothole. For instance, statistics released by the BBMP also mention that repairing a pothole once it has been formed completely is around three times the price of repairing it, if detected in its earlier stages. Through this paper we attempt to detect potholes and speed bumps on the fly, store their locations and use this data to alert/notify the user of an upcoming road anomaly so that precautionary measures in terms of reducing speed or taking an alternate route can be taken. To further enhance this system, the location of these road abnormalities can be made available to the concerned authorities for constant monitoring and maintenance.

B. U. Kempaiah (✉) · R. J. Mampilli · K. S. Goutham
Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_7

## 2 Related Work

Multiple approaches have been taken to tackle this problem. One of the very first was a study from Eriksson [6] that investigates an application for mobile sensing by detecting and reporting the surface conditions of roads. Multiple studies Silvaa [7] and Harikrishnan [8] took the data mining approach, to observe patterns in the data and extract the most important features and model them using a Machine learning approach. For detecting events and classification, K-Nearest Neighbour (K-NN) algorithm was used by Rajamohan [9], while Chen [10] uses Gaussian Mixture Model (GMM) and Improved Gaussian Mixture Model (i-GMM). Nature driven approaches such as genetic algorithms were also explored using images as a source of information to detect potholes and pavement distress were done by Celaya-Padilla [11], Bello-Salau [12], Salari [13]. Johnson [14] built a system known as MIROAD: A Mobile- Sensor- Platform for Intelligent Recognition of Aggressive Driving categorizes driving style into normal, aggressive and very aggressive. Multiple sensors are used (such as accelerometer, gyroscope, magnetometer, GPS, video) and data is coalesced into a single classifier based on Dynamic Time Warping (DTW) algorithm. Bhoraskar [15] uses accelerometer sensor data for detection of bumps and brakes. They also proposed an algorithm for reorienting the mobile phone device to align it with the vehicle axis, as the phone can be in any arbitrary location inside the vehicle. Numerous studies Sebestyen [16], Mohan [17], Forslof [18] have chosen to use smart phones and commodity sensors to collect the data and model the features using intelligent algorithms. Amita [19] is based on stereo-vision analysis of the road environments ahead of the vehicle. They have also developed 2 models based on deep-learning. The authors haven't considered different kinds of road anomalies and their main focus was towards the detection of potholes. Though the pothole detection is accurate when other anomalies such as speed bumps come into picture it would pose a greater challenge. Shah [20] classifies the road surface into speed-bump/potholes/normal roads based on image data using convolution neural networks. Furthermore, the YOLO algorithm was used to locate the pothole or speed bump. Although the study presented an 88.9% precision not much information is given on the source of the data used. The YOLO model can be easily avoided with the help of a gps system to avoid errors. Bansal [21] developed a machine learning based pothole detection system called DeepBus with the help of IoT sensors. The authors have used a completely ML based model for the detection of potholes. Since speed bumps are not considered, a naive ML model would show good results, but to differentiate between a speed bump and pothole, Recursive Neural Networks are preferred.

The solution proposed in the following section differs greatly from the above mentioned in terms of the hardware and the algorithm used. While a majority of the studies use smartphones, it adds uncertainty due to dissimilar sensor properties between various models of smartphones, as well as differences in vehicle size, weight, length, and suspension systems. In fact, different vehicles passing over a specific pothole would not generate an identical signal pattern. This study employs dedicated chips and sensors to record the signals. Machine Learning models such as Random Forests and Decision Trees are employed to detect anomalies, but they fundamentally differ from the solution mentioned in this study, since the former mentioned models do not take the sequence of sensor readings into account. The sensor values in succession reveal striking patterns that can be exploited to obtain significant improvement. A striking difference also lies in the region of study. A large proportion of the literature focuses on developing a solution in cities such as Boston, Portugal, Toronto that have a uniform road network with lower traffic densities with well-formed bumps and very few or no potholes at all. Very few studies are concerned with Indian cities. Additionally, this study also looks at employing the current state-of-the-art technology by storing data on the cloud and thus ensuring high availability, scalability and fault tolerance.

## 3 Proposed Approach

The study was conducted in Bangalore city where a good proportion of the roads are in unpleasant conditions, covered with potholes of varying sizes and speed bumps that do not conform to the regulations. The speed bumps in the city are unpainted and this reduces the effectiveness of previously proposed solutions that make use of image processing to tag the speed bumps. A variety of materials are used to build the roads ranging from asphalt to gravel and cement. Such diversities in the road anomalies greatly increase the difficulty of the task. The solution is broadly divided into 2 components, the Analysis Fig. 1 and Notification Fig. 2 modules.

### *3.1 Analysis Phase*

This phase begins with the collection of raw data from the accelerometer and GPS module. The raw data is pushed into the Cloud Storage at regular intervals to ensure a balance between the amount of data sent and the frequency of writes to the database. This raw data undergoes pre-processing as described in Sect. 4.2 and the pre-processed data is fed to the model described in Sect. 4.3 for classification. If the model identifies a road anomaly it is pushed onto the database to facilitate the notification phase. This entire process repeats in a loop to move forward in the Analysis phase. While the model remains trained initially on the available locations, new additions to the anomaly database is done when more a particular location is triggered more than 5 times. This ensures that locations that have temporary obstacles such as uneven surfaces are not flagged as speedbumps/potholes. An overview of this process can be seen in Fig. 1.

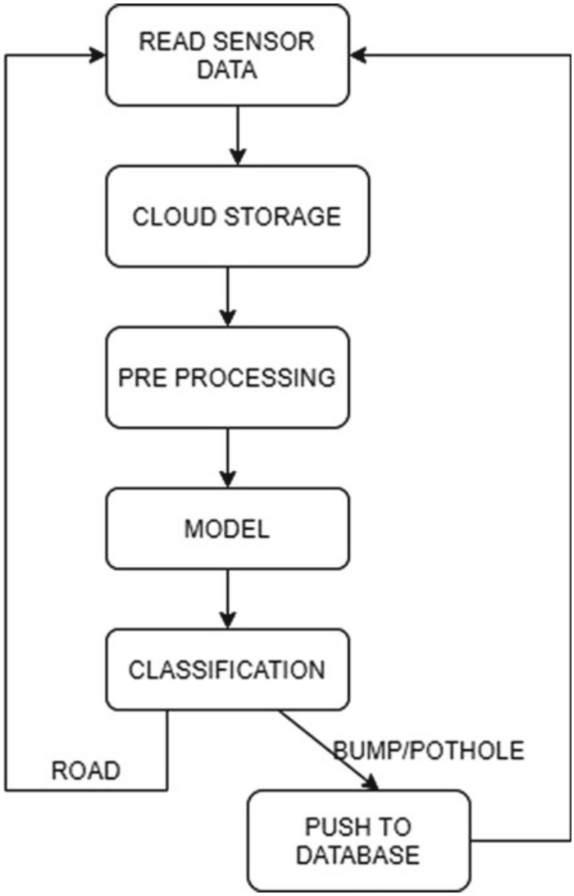



**Fig. 1** Analysis phase

### *3.2 Notification Phase*

The notification phase leverages the working of the Analysis phase**.** In this phase, the user's location is monitored constantly and if the user's location is in close proximity with any of the anomalies already flagged and present in the database, the user is notified. Else, the notification phase continues extracting the user's location. This is a continuous process that takes place throughout the journey of the commuter using the service.

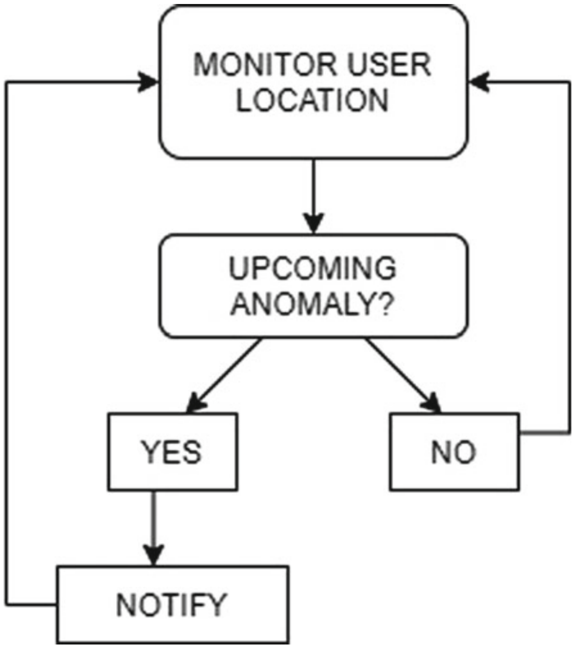



**Fig. 2** Notification phase

# 4 Methods

## *4.1 Data Collection*

Data was collected manually on the roads of Bangalore city. A speed of 10–20 kmph was maintained throughout the process. Custom hardware was employed to collect data to avoid the inconsistencies that come up when smartphones are used because the readings greatly depend on the orientation of the phone and its placement within the vehicle. The hardware module used in this study consists of a Raspberry pi with an ADXL 335 Triple Axis accelerometer and SKG13 GPS module. The hardware was mounted in the interior of the car. As the car starts to move, acceleration values along the $X$, $Y$ and $Z$ axis with the latitude and longitude from the GPS module were uploaded to a cloud storage powered by Google's Firebase. Sensor data was collected every 0.25 s and this raw data Fig. 3 was uploaded onto the cloud every three seconds to ensure the perfect balance between communication overhead and performance parameters. There were a total of 36,450 data points. Simultaneously, the co-passenger manually pinged a location which was then used to cross verify the model's prediction and to use it as a target column for training the model. The data was collected using 3 different vehicles varying in suspension quality to get an adequate representation of the real- world scenario of multiple users in various vehicles. Deploying the system on a cloud-based platform enables multi-user upload and download of data simultaneously.

| x | y | z | location | trial | direction | timestamp | lat | long | speed |
|---|---|---|---|---|---|---|---|---|---|
| -12 | -11 | -257 | CB | 28 | 0 | 02:35.7 | 13.06112 | 77.58513 | 5.021 |
| -8 | -4 | -241 | CB | 28 | 0 | 02:36.0 | 13.06112 | 77.58513 | 5.021 |
| -4 | -24 | -247 | CB | 28 | 0 | 02:36.2 | 13.06112 | 77.58513 | 5.021 |
| 5 | -38 | -251 | CB | 28 | 0 | 02:36.5 | 13.06109 | 77.58518 | 5.242 |
| -1 | -13 | -193 | CB | 28 | 0 | 02:36.7 | 13.06109 | 77.58518 | 5.242 |
| -26 | -6 | -251 | CB | 28 | 0 | 02:37.0 | 13.06109 | 77.58518 | 5.242 |
| 3 | -7 | -256 | CB | 28 | 0 | 02:37.2 | 13.06109 | 77.58518 | 5.242 |
| -4 | -23 | -249 | CB | 28 | 0 | 02:37.5 | 13.06107 | 77.58522 | 5.155 |

**Fig. 3** Raw data prior to adding the target column

| | | | | | | |
|---|---|---|---|---|---|---|
| -59 | 16 | -329 | 13.08132 | 77.66661 | 3.051 | 1 |
| -58 | 32 | -305 | 13.08133 | 77.66666 | 3.336 | 2 |
| -23 | -2 | -297 | 13.0825 | 77.6687 | 4.656 | 2 |
| -23 | 17 | -295 | 13.08232 | 77.67002 | 3.556 | 2 |
| 3 | -26 | -232 | 13.06103 | 77.58525 | 4.728 | 0 |
| -34 | -22 | -251 | 13.06046 | 77.58518 | 4.08 | 0 |
| -44 | 24 | -283 | 13.07993 | 77.66299 | 9.774 | 0 |
| -40 | -11 | -251 | 13.06033 | 77.58538 | 4.213 | 0 |

**Fig. 4** Data after pre-processing

## *4.2 Pre-processing of Data*

To ensure uniformity of data, the latitude and longitude values were brought down to 5 decimal places. Erroneous values were dropped from the dataset and the manually pinged locations were added onto the sensor data values to obtain complete data points Fig. 4. To further investigate, some preliminary tests were run. From Fig. 5 it is observed that the x and y-axis readings do not fluctuate to a great extent so as to be able to use them as principal axes for detecting the anomalies. On the other hand, z-axis is most indicative because when the vehicle moves over bumps and potholes, the vertical acceleration experiences significant changes. The trends for speedbumps and potholes are easy to identify in the case of z-axis because the variation experienced for road as seen by the orange line is very low. This characteristic of acceleration variation lies in the crux of the solution to the problem.

## *4.3 Detection Algorithm*

The data concerned in this study is a multivariate time series, hence a sequence modelling algorithm would produce appropriate results in classifying the road anomaly from the time series data. Traditional machine learning models such as

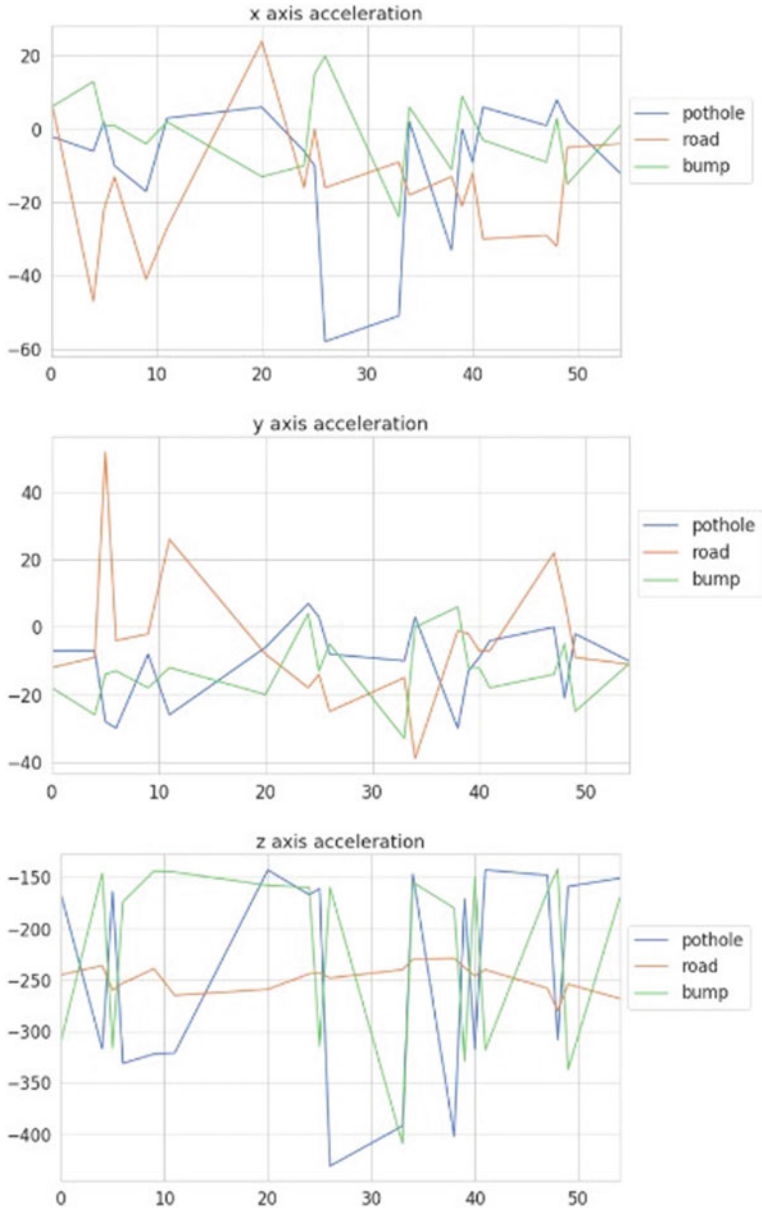



**Fig. 5** Variation of acceleration in the 3 axes

ANNs and SVMs work with one data point at a time which results in a loss of sequence and continuity in data resulting in failure of detecting bumps and potholes. This observation leads to Recurrent Neural Networks (RNN), which is a generalization of a Feedforward neural network that has an internal memory. After producing the output, it is copied and sent back into the recurrent network. The drawback of RNN is Gradient vanishing and exploding problems.

Long Short-Term Memory (LSTM) networks are a modified version of recurrent neural networks, which makes it easier to remember past data in memory. The vanishing gradient problem of RNN is resolved here. All recurrent neural networks have the form of a chain of repeating modules of neural networks. In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer Fig. 6. LSTMs also have this chain-like structure, but the repeating module has a different structure Fig. 7.

**Input gate**—discover which value from input should be used to modify the memory. Sigmoid function decides which values to let through 0,1. and tanh function gives weightage to the values which are passed deciding their level of importance ranging from −1 to 1.

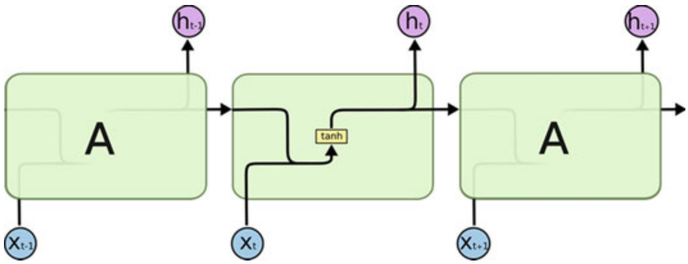



**Fig. 6** Schematic—recurrent neural network

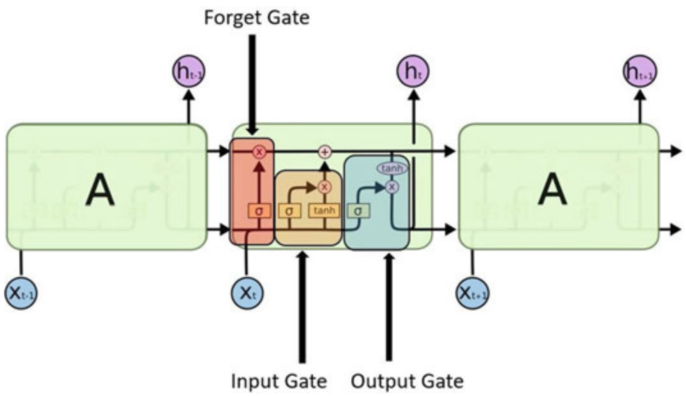



**Fig. 7** LSTM cell [22]

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{1}$$

$$C_t = \tan h(W_c[h_{t-1}, x_t]) + b_c) \tag{2}$$

| | |
|---|---|
| $x_t$ | input vector to the LSTM unit |
| $h_t$ | hidden state vector also known as output vector of the LSTM unit |
| $C_t$ | Cell state vector |
| $O_t$ | Output gate's activation vector |
| $i_t$ | input/update gate's activation vector |
| $f_t$ | forget gate's activation vector |
| *W, b* | Weight matrices and bias vector parameters which need to be learned during training |
| $\sigma$ | sigmoid function |
| tanh | hyperbolic tangent function. |

- **Forget gate**—discover what details to be discarded from the block. It is decided by the sigmoid function. It looks at the previous state ($h_{t-1}$) and the content input ($X_t$) and outputs a number between 0(omit this) and 1(keep this) for each number in the cell state $C_{t-1}$.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{3}$$

- **Output gate**—the input and the memory of the block is used to decide the output. Sigmoid function decides which values to let through and tanh function gives weightage to the values which are passed deciding their level of importance ranging from-1 to 1 and multiplied with output of Sigmoid.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{4}$$

$$h_t = o_t \times \tanh(C_t) \tag{5}$$

Various parameters of the model can be tuned such as the number of hidden layers, the time step, the activation function, the loss function and the optimizer function to name a few. The algorithm acts upon the raw data pushed onto the cloud and moves a location to the confirmed bucket only if the location has been pinged by the algorithm more than three times by various users.

## *4.4 Notification*

Alerting a user of an upcoming anomaly is done by constantly monitoring their location and analyzing whether the user is in close proximity to any of the already flagged bumps or potholes. A web-based interface has been built which is currently

**Fig. 8** User interface for notification

in its beta version, giving a visual appeal about the presence of speed bumps and potholes by suitable indication marks on the map. The interface provided to the user appears as shown in Fig. 8. A voice message informing the commuter “Speedbump ahead” is also present. While this could appear to be a distraction to some users, future enhancements can be made where a user can customize notifications received. Another application is to notify the concerned authorities with the locations of the road anomalies at regular intervals.

## 5 Results

It is observed that employing intelligent algorithms with sufficient amounts of data provides feasible results. The confusion matrices tabulated Tables 1 and 2 depict the performance of the system in the testing phase. The time step considered in this test is eight which means each sequence will contain eight data points. This is

**Table 1** Results of the 3-class classifier

| Predicted → Road | (0) Road | (1) Speedbump | (2) Pothole true |
|---|---|---|---|
| (0) Road | 524 | 14 | 20 |
| (1) Speedbump | 40 | 330 | 80 |
| (2) Pothole | 36 | 60 | 163 |

Average Accuracy = 0.8
Precision of Speed Bump = 0.81

**Table 2** Results of the 2-class classifier

| Predicted → True | (0) Road | (1) Road Anomaly |
|---|---|---|
| (0) Road | 264 | 27 |
| (1) Road Anomaly | 20 | 333 |

Average Accuracy = 0.93
Precision of Road Anomaly = 0.925

approximately data collected over two seconds. The Activation Function used at the hidden layer is ReLu. Two LSTM layers were used each having 64 hidden nodes. Cross Entropy Loss was used as the loss function and Adam as the optimizer function. The data was divided in the 80–20 ratio with the latter being the test size.

The results are represented in the form of a confusion matrix as seen in Tables 1 and 2. Each row of the matrix represents the instances in a predicted class while each column represents the instances in the actual class.

As seen in Table 1, the model is performing satisfactorily with an accuracy of 80%. This implies the model is able to classify a data point correctly 80 out of 100 times.

To get an idea of how the classifier performs when there exists no distinction amongst road anomalies Table 2, the accuracy increases significantly indicating that the task of differentiating a smooth road from rough one, is comparatively easier. In a two-class classifier, the model is able to correctly classify 93 out of every 100 data points.

The classifier performs satisfactorily and such a system would definitely provide efficacious results, when provided with adequate computation and storage facilities. On average, it takes 3.02 seconds to upload batches of data points onto the cloud from the hardware module and 3.4 seconds for the algorithm to output a class.

## 6 Future Work

Due to limited resources, only a total of 20 km around Bangalore was used to collect data. This can be extended to ensure that the entire road network is mapped. Since road anomalies are not constrained to just speed bumps and potholes, other anomalies such as railway crossings, rough patches, gravel etc. can be included for classification. Classifying various types of anomalies poses a greater challenge due to the large number of classes. Using multiple inputs alongside analog sensors such as imaging data etc. will tremendously add to the value of the system. Employing an ensemble of learning algorithms that simultaneously work on multi-input data can improve the efficiency of the solution manifold. In a real-time system involving a large network of users, multiple external parameters come into play which have to be looked into for further development. These include the latency involved in pushing the data onto the cloud and for the application to constantly receive data from the cloud to notify users suitably. All these factors can be considered for further enhancement of the system.

## References

1. Bruhat Bengaluru Mahanagara Palike (BBMP) (2018) engineers have drawn up a staggering Rs 1719-crore plan to fill potholes as and when they occur over the next two years. https://timesofindia.indiatimes.com/city/bengaluru/bbmp-plots-rs-1700-crore-plan-to-fix-potholes/articleshow/66079038.cms
2. Potholes are one of the factors that contribute to traffic congestion. https://economictimes.indiatimes.com/news/politics-and-nation/its-a-hell-of-a-ride-on-bangalore-roads/articleshow/70886816.cms?from=mdr
3. 25% of roads in Bengaluru have potholes according to BBMP. https://timesofindia.indiatimes.com/city/bengaluru/25-of-bengaluru-roads-have-potholes-concedes-bbmp/articleshow/73209775.cms
4. 175 potholesin a stretch of 800 metres. https://bangaloremirror.indiatimes.com/bangalore/civic/bengaluru-800-metre-stretch-175-potholes/article show/71250667.cms
5. Speed Bumps that do not fall under the regulations are sources to many accidents. https://www.deccanherald.com/city/speed-kills-so-do-speed-698933.html
6. Eriksson J, Girod L, Hull B, Newton R, Madden S, Balakrishnan H (2008) The pothole patrol: using a mobile sensor network for road surface monitoring. https://doi.org/10.1145/1378600.1378605
7. Silvaa N, Soaresa J, Shaha V, Santosa MY, Rodriguesa H (2017) Anomaly detection in roads using a data mining approach. https://doi.org/10.1016/j.procs.2017.11.056
8. Harikrishnan PM, Gopi VP (2017) Vehicle vibration signal processing for road surface monitoring. IEEE Sensors J 17(16):5192–5197. https://doi.org/10.1109/JSEN.2017.2719865
9. Rajamohan D, Gannu B, Rajan K (2015) MAARGHA: a prototype system for road condition and surface type estimation by fusing multi-sensor data. ISPRS Int J Geo-Info 4:1225–1245. https://doi.org/10.3390/ijgi4031225
10. Chen K, Tan G, Lu M et al (2016) CRSM: a practical crowdsourcing-based road surface monitoring system. Wireless Netw 22:765–779. https://doi.org/10.1007/s11276-015-0996-y
11. Celaya-Padilla JM, Galván-Tejada CE, López-Monteagudo FE, Alonso-González O, Moreno-Báez A, Martínez-Torteya A, Galván- Tejada JI, Arceo-Olague JG, Luna-García H, Gamboa-Rosales H (2018) Speed bump detection using accelerometric features: a genetic algorithm approach. Sensors 18:443. https://doi.org/10.3390/s18020443
12. Bello-Salau H, Aibinu A, Onwuka E, Dukiya J, Onumanyi A (2014) Image processing techniques for automated road defect detection: a survey. In: Proceedings of the 2014 11th international conference on electronics, computer and computation (ICECCO), Abuja, Nigeria, 29 September–1 October 2014; pp 1–4. https://doi.org/10.1109/ICECCO.2014.6997556

13. Salari E, Yu X (2011) Pavement distress detection and classification using a Genetic Algorithm. In: Proceedings of the 2011 applied imagery pattern recognition workshop (AIPR), Washington, DC, USA, 11–13 October 2011; pp 1–5. https://doi.org/10.1109/AIPR.2011.6176378
14. Johnson DA, Trivedi MM (2011) Driving style recognition using a smart- phone as a sensor platform. In: IEEE 14th international conference on intelligent transportation system, October (2011). https://doi.org/10.1109/ITSC.2011.6083078
15. Bhoraskar R, Vankadhara N, Raman B, Kulkarni P (2012) Wolverine: traffic and road condition estimation using smartphone sensors. In: Fourth international conference on communication systems and networks (COM- SNETS). IEEE (January 2012). https://doi.org/10.1109/COMSNETS.2012.6151382
16. Sebestyen G, Muresan D, Hangan A (2015) Road quality evaluation with mobile devices. In: Proceedings of the 2015 16th international carpathian control conference (ICCC), Szilvasvarad, pp 458–464. https://doi.org/10.1109/CarpathianCC.2015.7145123
17. Prashanth M, Venkata N, Ramjee PR (2008) Nericell: rich monitoring of road and traffic conditions using mobile sensors. Bangalore, Microsoft Research India. https://doi.org/10.1145/1460412.1460444
18. Forslof L (2012) Roadroid: smartphone road quality monitoring. In: Road-roid, AB Sweden, Egnahemesgatan, October 2012
19. Dhiman A, Klette R (2020) Pothole detection using computer vision and learning. IEEE Trans Intell Transp Syst 21(8):3536–3550. https://doi.org/10.1109/TITS.2019.2931297
20. Shah S, Deshmukh C (2019) Pothole and bump detection using convolution neural networks. In: 2019 IEEE transportation electrification conference (ITEC-India), Bengaluru, India, pp 1−4. https://doi.org/10.1109/ITEC-India48457.2019.ITECINDIA2019-186
21. Bansal K, Mittal K, Ahuja G, Singh A, Gill SS (2020) DeepBus: Machine learning based real time pothole detection system for smart transportation using IoT. Internet Technol Lett. https://doi.org/10.1002/itl2.156
22. Figure for LSTM cell https://colah.github.io/posts/2015-08-Understanding-LSTMs/#:~:text=Long%20Short%20Term%20Memory%20networks,many%20people%20in%20following%20work

# Decision Tree Based Crop Yield Prediction Using Agro-climatic Parameters

**K. Aditya Shastry, H. A. Sanjay, and M. C. Sajini**

## 1 Introduction

Currently, DM is being utilized for performing tasks such as prediction, association analysis, clustering, and classification in several applications. Numerous DM software's are also available [1]. However, the application of DM in the domain of agriculture, particularly in developing countries like India, is limited [2]. Vast amounts of data are being generated with respect to crops, soils, weather, etc. The analysis of such agricultural datasets through DM techniques can beneficially impact the way traditional farming is done [3]. Early forecasting of crop yield is a significant aspect of agriculture. Though research is going on in developed countries regarding crop yield prediction (CYP), its usage in developing countries is less. Early prediction of yields of significant crops such as wheat, soybean, rice, etc. can lead to reasonable profits for the farmers [4].

Keeping these points, in mind this research aims in doing an analysis of the soybean crop dataset and predict its yield. The soybean dataset consisting of rainfall (in mm), evaporation (in mm), temperature (in degrees), relative humidity (in percentage), and historic soybean yield (in tons per hectare) was analyzed. The yield of soybean was predicted as high (H) or low (L) using the ID3 algorithm. The result was compared with the NB classifier with respect to performance metrics like True Positive (TP) rate, False Positive (FP) rate, F-Score, Recall, Precision, and receiver operating characteristic curve (ROC) Curve. Results demonstrated that the ID3 algorithm provided an accurate estimate of the yield than the NB classifier. The motivation in taking up this work was the lack of efficient application of DM techniques in the

K. Aditya Shastry (✉) · H. A. Sanjay
Nitte Meenakshi Institute of Technology, Bengaluru 560064, India
e-mail: adityashastry.k@nmit.ac.in

M. C. Sajini
Tata Consultancy Services, Pune 411057, India

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_8

prediction of crop yield in India. Some of the challenges faced in implementing this work are highlighted below:

- Noisy data: Data containing random components such as corrupted, missing, and inconsistent values.
- Timeliness: The data set starts aging as soon as it is collected
- Relevance: The data set must consist of relevant information or attributes.

The structuring of the paper is as follows. Section 2 reviews the background work in the domain of CYP using DM techniques. Section 3 describes the specifics of the proposed work. Section 4 expresses the experimental details and results attained along with the comparison of results. The paper ends with the conclusion and future work.

## 2 Related Work

In this section, we survey some of the recent works in the domain of CYP using DM algorithms.

In [5], the authors estimated the wheat yield by utilizing the crop simulation module called CERES-Wheat in DSSAT environment in 6 regions of India. They followed a 3-step procedure, viz (i) forecasting of trend-based yields using technology, (ii) the yield variability due to weather was quantified using crop simulation model (CSM), (iii) The prediction of final yield was performed by combining steps (i) and (ii). The authors were able to assess the variability in the yields of wheat in real time. The WEKA model was also utilized for building the process model. The authors planned to instill the domain model learned by the DM algorithm in software applications.

The work [6], deals with the application of DM techniques from WEKA tool to predict the yields of maize, soybean, and sugar beet. Authors made use of climatic attributes such as temperature, precipitation, etc. They concluded that the M5P model tree produced better regression rules. Comparison with other methods was not performed.

Four regression strategies viz. Regression trees, Multilayer Perceptron (MLP), Support Vector Regression (SVR), Regression Tree, and Radial Basis Function (RBF) models were utilized in [7] to forecast the yield of winter wheat using fertilization parameters. Authors concluded that the SVR model performed better prediction of wheat yield. In [8], the authors demonstrated the effectiveness of association rule mining using a low support threshold on data set that was dense in nature. Apriori algorithm was utilized for association analysis in order to forecast the chances of disease affecting the area. This work generated 3 diverse rule sets on the agricultural dataset by utilizing the Apriori algorithm. The dependencies or the relationships among the features were found.

It is evident from the above survey, that in developing countries the DM methods are not used extensively in the agricultural domain. In most cases, regression-based tools are utilized for identifying the impact of agricultural parameters on the yield of a crop. The major limitations of the above works were that either they were performed on smaller agricultural areas or the impact of the input features were not studied. In view of this, our work proposes to develop a DM based application for predicting crop yield using the critical agro-climatic factors.

## 3 Proposed Work

This section describes the proposed work of predicting the soybean yield using ID3 algorithm. Figure 1 depicts the methodology followed for the current work. The general modules used in our work are illustrated in the following figure.

The dataset can be read from a text file or from MySql database. The user enters values for the input attributes. When the user presses the submit button, the corresponding data gets stored in the database. This dataset can be used as a training dataset for training the ID3 algorithm. The ID3 algorithm is then applied to the dataset. The Entropy ($E$) and information gain (IG) for each attribute of the dataset is computed. The Entropy ($E$) is calculated using the Eq. 1.

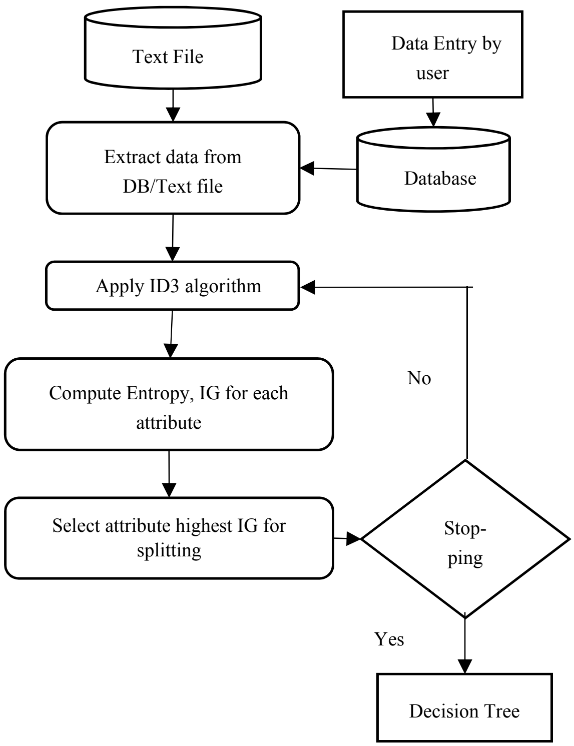



**Fig. 1** Proposed methodology for forecasting crop yield using ID3 algorithm

$$E = -\sum_{i}^{C} p_i \log_2 p_i \tag{1}$$

where, $E$ is entropy, $p_i$ denotes probability of randomly picking an element of class $i$, and $C$ is the number of classes. The IG is computed using Eq. 2.

$$IG(T, a) = E(T) - E(T|a) \tag{2}$$

where IG is the Information Gain, $T$ is the Training Data, '$a$' is the attribute, and '$E$' is the Entropy. $E(T|a)$ signifies the conditional entropy of $T$ give value of '$a$'. The attribute with the highest IG is selected for splitting. The process is repeated for all attributes until no attribute is further left for partition (i.e., stopping criteria is true). Finally, the Decision tree is generated. The results of the decision tree are compared with the naïve Bayes classifier.

## 4 Experimental Setup and Results

In this section, we discuss the setup and results of this work. The experiments were carried out on Windows 10 operating system with 2 TB of HDD, and 6 GB of RAM. Java jdk 1.6 was used as the front-end, netbeans-7.2.1 as the IDE, and MySql-5.5.28 as the backend. The data set considered for the experiment was the soybean crop data set from Bhopal District of Maharastra, India [9]. The dataset consisted of the attributes such as: Rainfall (in mm), Evaporation, Temperature (in degrees), Relative Humidity (in percentage), and historic Soybean Yield (in tons per hectare). The above


**Fig. 2** Login screen

data was collected from 1980 to 2019 consisting of 150 records. Figure 2 Shows the login screen through which the user logs in.

If the user inputs the correct username and password, then he/she goes to the next screen shown in Fig. 3.

In Fig. 3, when the user enters the rainfall, evaporation, temperature, and relative humidity, the corresponding soybean yield is predicted using the ID3 algorithm. Figure 4 depicts the screenshot that shows how the results are visualized by the end user.

As illustrated in Fig. 4, the yield of soybean is predicted in the form of low and high yields. Figure 5 depicts the dataset used for our experiments.

Figure 6 depicts the resulting DT constructed by the ID3 algorithm.

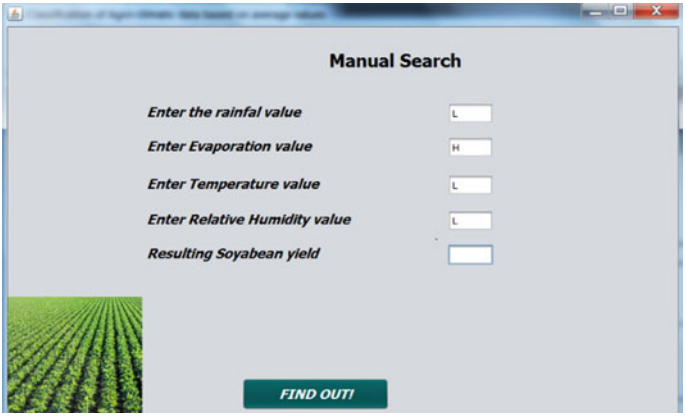



**Fig. 3** Manual search

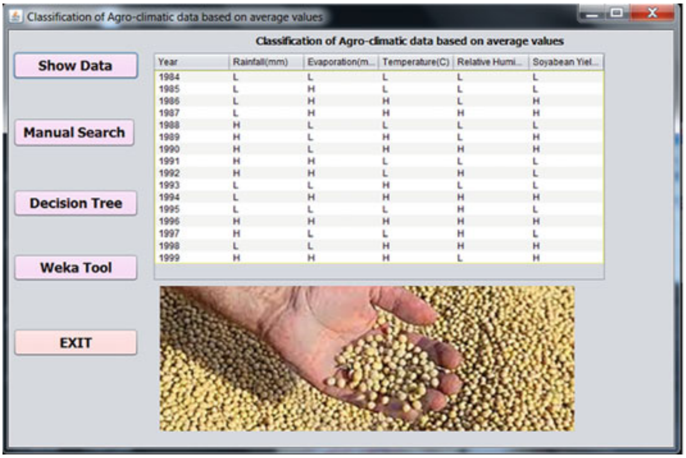
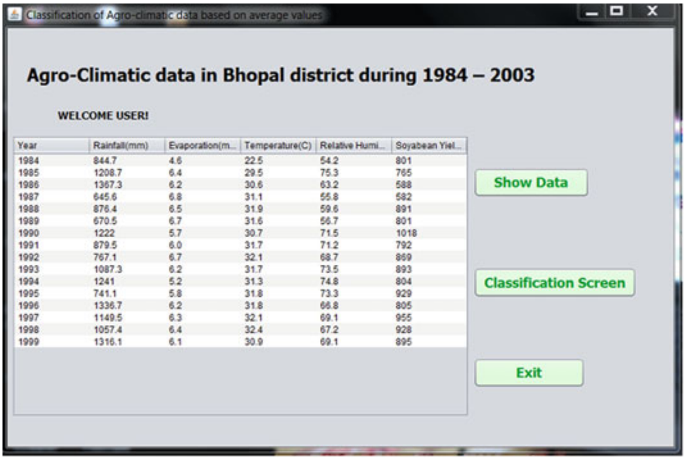



**Fig. 4** Dataset prediction based on threshold values




| Year | Rainfall(mm) | Evaporation(m... | Temperature(C) | Relative Humi... | Soyabean Yiel... |
|---|---|---|---|---|---|
| 1984 | 844.7 | 4.6 | 22.5 | 54.2 | 801 |
| 1985 | 1208.7 | 6.4 | 29.5 | 75.3 | 765 |
| 1986 | 1367.3 | 6.2 | 30.6 | 63.2 | 588 |
| 1987 | 645.6 | 6.8 | 31.1 | 55.8 | 582 |
| 1988 | 876.4 | 6.5 | 31.9 | 59.6 | 891 |
| 1989 | 670.5 | 6.7 | 31.6 | 56.7 | 801 |
| 1990 | 1222 | 5.7 | 30.7 | 71.5 | 1018 |
| 1991 | 879.5 | 6.0 | 31.7 | 71.2 | 792 |
| 1992 | 767.1 | 6.7 | 32.1 | 68.7 | 869 |
| 1993 | 1087.3 | 6.2 | 31.7 | 73.5 | 893 |
| 1994 | 1241 | 5.2 | 31.3 | 74.8 | 804 |
| 1995 | 741.1 | 5.8 | 31.8 | 73.3 | 929 |
| 1996 | 1336.7 | 6.2 | 31.8 | 66.8 | 805 |
| 1997 | 1149.5 | 6.3 | 32.1 | 69.1 | 955 |
| 1998 | 1057.4 | 6.4 | 32.4 | 67.2 | 928 |
| 1999 | 1316.1 | 6.1 | 30.9 | 69.1 | 895 |

**Fig. 5** Numeric dataset for soybean productivity

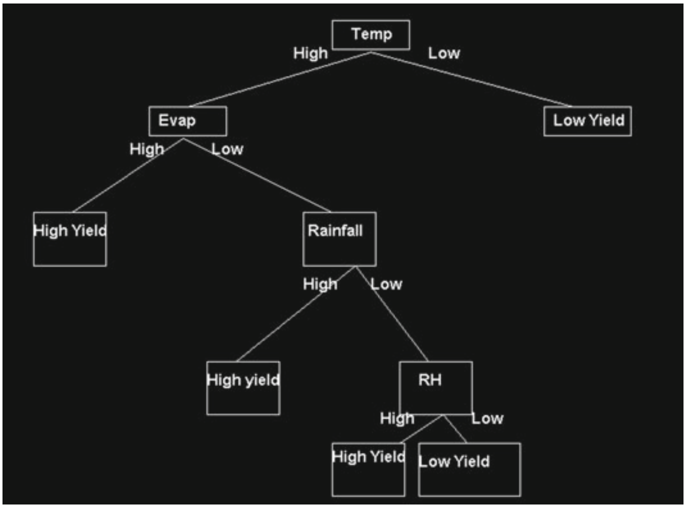



**Fig. 6** Numeric dataset for soybean productivity

As illustrated in Fig. 6, the Temp attribute was chosen as the root node, since it had the highest IG. The leaf nodes represented the High and Low yields of soybean. The attribute with the highest information gain was selected for splitting. The results of the ID3 were compared with the NB classifier. Figure 7 illustrates the comparison of ID3 with NB classifier using performance metrics like Accuracy, Precision, Recall, F-Measure, and ROC.

**Fig. 7** Result comparison

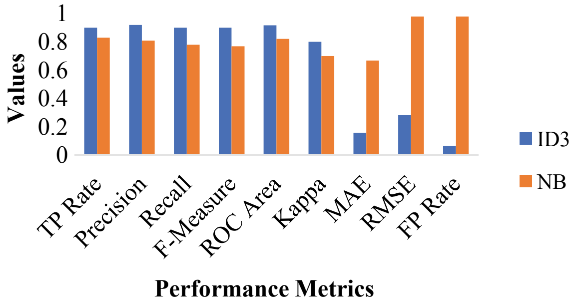



**Table 1** Values of performance metrics of ID3 and NB classifiers

| Heading level | ID3 | NB |
|---|---|---|
| TP Rate | 0.9 | 0.83 |
| Precision | 0.92 | 0.81 |
| Recall | 0.9 | 0.78 |
| F-Measure | 0.901 | 0.77 |
| ROC area | 0.917 | 0.821 |
| Kappa | 0.8 | 0.7 |
| MAE | 0.16 | 0.67 |
| RMSE | 0.2828 | 0.98 |
| FP Rate | 0.067 | 0.98 |

As can be observed from the above fig, the ID3 algorithm performed with higher TP rate, precision, recall, F-Measure, ROC area, Kappa Statistic than the NB classifier. Similarly, it had lesser RMSE, MAE, and FP rate than the NB classifier. The exact values of the performance metrics are tabulated in Table 1.

## 5 Conclusion and Future Work

This work applies the decision tree algorithm for the prediction of crop yield. In this work, we have attempted to stress the idea that previous historic agricultural data along with the related climatic factors can be successfully used to predict the future yield of crop. Specifically, the soybean crop is considered for experimentation. This work helps the farmers to have a better idea of the climatic factors that affect crop production. The DT shows the graph for our dataset for which the attribute which has the highest IG is taken as the head node. The results demonstrated that the ID3 algorithm performed better than the NB classifier with respect to several performance metrics.

In our work, the experiments have been carried out for predicting the soybean yield on a single processor system using the ID3 algorithm. In the future, the system can be modified to work on large datasets using a distributed approach. Additionally, the methods can be more generalized by interested researchers using an alternative approach such as bottom-up instead of the top-down approach used in this work. The technique used in this paper can also be applied in different domains other than agriculture. Sophisticated pruning techniques may also be applied for algorithms related to decision trees. This can reduce the complexity of the algorithm significantly. In the future, experiments may be conducted on several other crop datasets instead of focusing on one specific crop dataset as was done in this work. Exhaustive comparisons with other well-established algorithms may be performed in the future.

## References

1. Agarwal S (2013) Data mining: data mining concepts and techniques. In: International conference on machine intelligence and research advancement, IEEE, Katra, India, pp 203–207. https://doi.org/10.1109/ICMIRA.2013.45
2. Gandhi N, Armstrong LJ, Nandawadeka M (2017) Application of data mining techniques for predicting rice crop yield in semi-arid climatic zone of India. In: Technological innovations in ICT for agriculture and rural development (TIAR), IEEE, Chennai, India, pp 116–120. https://doi.org/10.1109/TIAR.2017.8273697
3. Vanitha CN, Archana N, Sowmiya R (2019) Agriculture analysis using data mining and machine learning techniques. In: 5th International conference on advanced computing and communication systems (ICACCS), IEEE, Coimbatore, India, pp 984–990. https://doi.org/10.1109/ICACCS.2019.8728382
4. Chlingaryan A, Sukkarieh S, Whelan B (2018) Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Comput Electron Agricul 151:61–69. https://doi.org/10.1016/j.compag.2018.05.012
5. Veenadhari S, Bharat M, Singh Cd (2011) Soybean productivity modelling using decision tree algorithms. Int J Comput Appl 27(7):11–15. https://doi.org/10.5120/3314-4549
6. Marinković B, Crnobarac J, Brdar S, Antić B, Jacimovic G, Crnojević V (2009) Data mining approach for predictive modeling of agricultural yield data. In: BioSense 2009 sensing technology in agriculture, forestry and environment workshop, Novi Sad, Serbia, pp 1–5
7. Georg R (2009) Data mining of agricultural yield data: a comparison of regression models. In: Perner P (eds) Advances in data mining. applications and theoretical aspects, ICDM, Lecture notes in computer science, vol 5633, pp 24–37. https://doi.org/10.1007/978-3-642-03067-3_3
8. Smitha T, Sundaram V (2012) Association models for prediction with Apriori concept. Int J Adv Eng Technol 5(1):354–360
9. Indian Agricultural statistics Homepage (2017). http://agricoop.nic.in/statistics/state-level. Last Accessed 17 Apr 2017

# Regression Based Data Pre-processing Technique for Predicting Missing Values

**K. Aditya Shastry, H. A. Sanjay, and M. S. Praveen**

## 1 Introduction

DM has developed as one of the prominent research domains in the recent years for extracting meaningful information / knowledge from large amounts of data. This knowledge can be easily comprehended by humans [1]. The raw data is transformed into a format that is understandable in data pre-processing. Often, we observe that the real-world data is inconsistent, incomplete, prone to typing errors. These issues can be resolved using data pre-processing. It prepares the raw data to be processed further. Cleaning, transformation, integration, reduction and discretization of data form the major steps in data pre-processing [2].

Techniques for data pre-processing are required since data in the real world is noisy, dirty and incomplete. Incomplete data means that certain values of attributes may be missing or only aggregate of values may be present. Certain data in specific instances or records may be missing. The incomplete data can arise when data being collected from the source has "Not Applicable (NA)" values. Incorrect entries by humans may cause noisy data to be generated. Different data sources may cause inconsistencies in data. Records containing duplicate values need to be removed. Before any data mining algorithm is applied, the data needs to be cleaned by performing data pre-processing [3].

Some of the issues usually encountered in the data pre-processing are [3]:

K. Aditya Shastry (✉) · H. A. Sanjay
Nitte Meenakshi Institute of Technology, Bengaluru 560064, India
e-mail: adityashastry.k@nmit.ac.in

H. A. Sanjay
e-mail: sanjay.ha@nmit.ac.in

M. S. Praveen
Tata Consultancy Services, Pune 411057, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_9

- Noisy Data: Noise is the random component that occurs due to error in measurements. There are two types of noises. One type of noise are the corrupted values. These values are altered in the training set leading to conflicts in the tuples of the data. These values can cause extremities and hence are considered as noise. They can be ignored during the DM process. The 2nd type of noise are the missing attribute values. The training set may comprise of 1 or more missing values in certain attributes or records. They may occur since the data may not be relevant for a attribute or is not recorded while data is being collected or may have been ignored due to privacy concerns. Missing values in data may lead to incorrect predictions by the data mining algorithm. Hence, they must remove, or a suitable value must be put in place of the missing values.
- Difficult Training Set: The training data may be difficult to interpret and comprehend making the whole process of mining the data time consuming
- Huge Database: Datasets that are huge in nature need to be pre-processed appropriately in order to reduce the complexity and time [3].

The chief objective of this work is the prediction of missing values in the given raw data using multiple linear regression (MLR). This prediction is performed by utilizing specific selected attributes of the raw data. The best prediction model is chosen based on performance evaluation of the models using the performance metrics like Root Mean Squared Error (RMSE) and R-Squared ($R^2$). Specifically, this paper discusses the reconstruction of incomplete data (missing values) by the technique of statistical reconstruction the data utilizing MLR. The missing values are predicted based on the correlation of the attributes that help to generate data sets that are complete in nature.

## 2 Related Work

In this section, we deliberate some of the relevant works in the domain of data pre-processing for missing value prediction.

The work [4] focuses on the development of ARLSimpute which represents the autoregressive model for predicting missing values. The output of the data pre-processing is provided as input to the prediction techniques namely quadratic & linear prediction. These methods are utilized to forecast the future values based on the past values.

Authors in [5] devised a technique for handling missing values on NCDC weather dataset. They estimated the prediction error of 5 techniques viz. random forest, kernel ridge, K-NN & SVM imputation. The predicted missing values were compared with the actual value for each method.

Kotteti et al. [6], imputed missing categorical values using most frequent value in column and for numerical features mean of the column was utilized. They demonstrated that Multilayer Perceptron (MLP) with the proposed data pre-processed method performed better.

Zeng et al. [7], evaluate the performance of several imputation methods. The authors, simulate the missing data set by completely deleting some data from the complete data set, and use the Euclidean distance KNN, the correlation coefficient KNN and the mean to fill several algorithms to estimate the exact data and compare the accuracy of different algorithm estimation.

## 3 Methodology

The general modules used in this work are illustrated in the Fig. 1. In the first module we read the dataset as the EXCEL and WORD file. The data must be pre-processed. The next module pre-processes the dataset into sample file, from which the required dataset is divided into two parts, called training data set and test dataset. Sampling is done which splits into training dataset and test dataset. Then again pre-processing is done for split dataset. The training dataset has 70% and test dataset has 30%. Many regression models such as LR, IR, PQR, QR, and PR are applied for prediction of missing values.

Initially, the regression models are applied on the training dataset. Of all these models whichever gives the best accuracy is chosen to be the model for the test dataset. The new data is constructed which is free of errors and missing values. The noisy, irrelevant and redundant information are eliminated. After data pre-processing, the knowledge discovery needs to be done which is a much more difficult task. However, data pre-processing forms a critical step in knowledge discovery as inaccurate data can lead to wrong predictions. The Data pre-processing constitutes transformation,

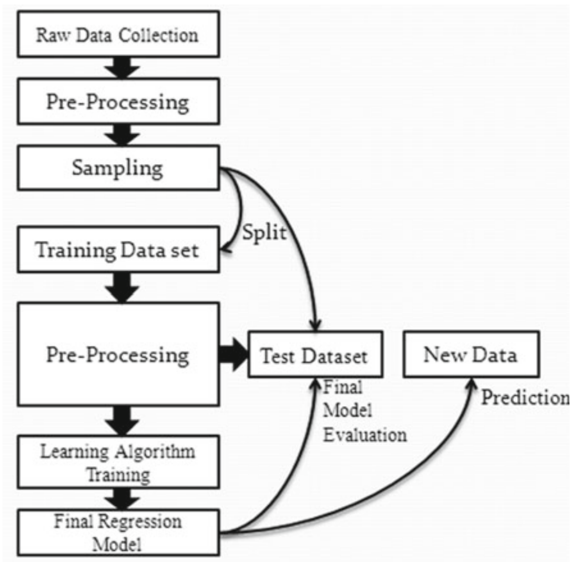



**Fig. 1** Proposed methodology

normalization, feature extraction & selection. The outcome of the data pre-processing is the complete pre-processed dataset.

In our work, following regression models were applied to predict the missing values.

- LR Model: It contains independent (explanatory) variable, Xi, for $i = 1\ldots$ n variables. The parameters of regression and the dependent variables are linear in nature. The estimation of the dependent variable is done. Equation 1 represents the LR model.

$$Y = a + \sum_{i=1}^{n} bX_i + e^i \tag{1}$$

  where,

  - '$a$'➔ intercept
  - '$n$'➔ number of features,
  - '$b$'➔ slope of the regression line.
  - '$e_i$'➔ random error term that is assumed to be uncorrelated [8].

- QR Model: It models the relationship among the sets of variables. Prediction about the data is the outcome of the QR model. The QR model is shown in Eq. (2):

$$y = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i + c \tag{2}$$

where,

- '$a$', '$b$' and '$c$' ➔ coefficients,
- '$y$' ➔ dependent variable,
- '$x$' ➔ explanatory variable and
- '$n$' ➔ number of features [9].

- PQR Model: This model is like the QR model except that it contains the squared terms. The PQR model does not contain terms that are raised to the power of 1. It is usually represented using the general format $kx^2 + m = 0$. For example, $x^2 = 38$ denotes a PQR model in which $k = 1$, and $c = -48$ [10].
- IR Model: An IR model consists of variables that are in constant interaction with one another. The dependent variables are interacting with each other [10].
- PR Model: It represents regression analysis which models the relationship among independent variable '$x$' and the dependent variable '$y$' as $n$th degree polynomial in $x$. It fits a relationship between $x$ and $y$ which is nonlinear in nature [10].

## 4 Experimental Setup and Results

Experiments were carried out on stock dataset obtained from KEEL repository [11]. Matlab R2016a was used as the scripting language. The dataset comprised of day wise prices of stocks of 10 aerospace organizations during the period 1998 Jan through 1991 October. The task was to forecast the stock price of the 10th company based on the prices of the remaining companies. The stock dataset consisted of 950 records with 9 input attributes.

Comparison was made between linear, pure-quadratic, interactions, quadratic, and polynomial regression models. Results demonstrated that the polynomial regression model performed with higher accuracy in terms of R2 and lower error rate in terms of RMSE. Figure 2 shows the comparison of the PCA and the regression models for forecasting missing values on the stock dataset.

As can be observed from Fig. 2, the PR model predicted the missing values in the stock dataset more accurately than the LR, PQR, IR and QR models.

Table 1 shows the sample prediction done by the PR model on the stock dataset.

Table 2 depicts the comparison of the regression models with respect to RMSE and $R^2$.

We tested with diverse polynomial models by varying their degrees and found that the following PR model gave the best result.

$$y = x_1^3 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 \tag{3}$$

where,

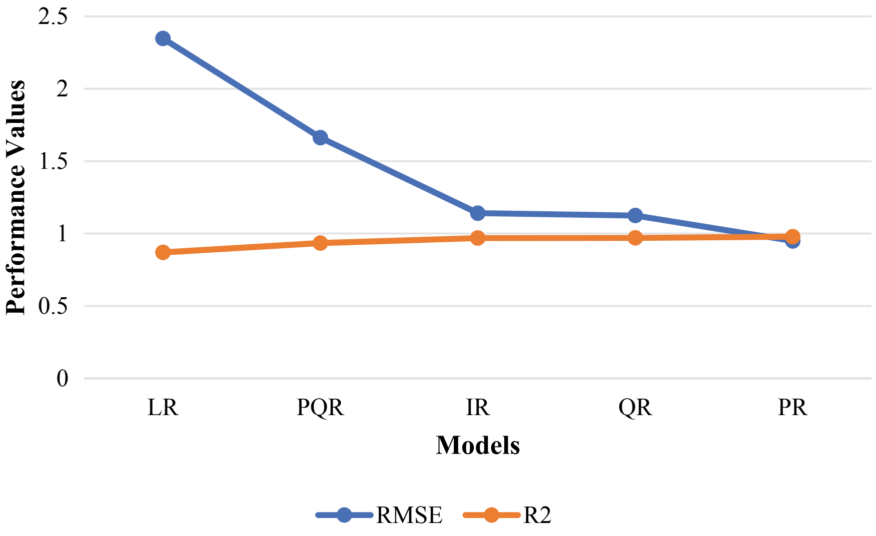



**Fig. 2** Model comparison for missing value prediction

**Table 1** Sample predictions of PR model on stock dataset

| Actual value | Predicted value | Error rate (%) | Accuracy (%) |
|---|---|---|---|
| 51 | 40.4581 | 10.54 | 89.46 |
| 39.875 | 35.7314 | 4.14 | 95.85 |
| 38.625 | 31.8732 | 6.75 | 93.25 |
| 38.375 | 46.5222 | 8.17 | 91.83 |
| 59.875 | 21.2866 | 38.58 | 61.4 |
| 46.25 | 50.6010 | 4.35 | 95.65 |
| 45.625 | 59.8302 | 14.2 | 85.8 |
| 45.25 | 62.1199 | 16.86 | 81.14 |
| 44.25 | 51.8949 | 7.64 | 92.36 |
| 37.375 | 35.2795 | 2.09 | 97.91 |

**Table 2** Comparison of regression models

| Models | RMSE | $R^2$ |
|---|---|---|
| LR | 2.3476 | 0.8703 |
| PQR | 1.6625 | 0.9348 |
| IR | 1.1409 | 0.9693 |
| QR | 1.1249 | 0.9702 |
| PR | 0.9496 | 0.9786 |

- $y$ signifies the target attribute denoting the stock price of the 10$^{\text{th}}$ company.
- $x_1$ to $x_9$ represent the input attributes that denote the stock prices of 9 aerospace companies.

Equation 2 represents the LR model:

$$y = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 \tag{4}$$

Equation 3 signifies the QR model

$$\begin{aligned} y &= x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 \\ &\quad + x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2 + x_8^2 + x_9^2 \end{aligned} \tag{5}$$

Equation 4 denotes the PQR model

$$y = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2 + x_8^2 + x_9^2 \tag{6}$$

Equation 5 represents the IR model

$$\begin{aligned} y &= x_1 * x_2 + x_2 * x_3 + x_3 * x_4 + x_4 * x_5 + x_5 * x_6 \\ &\quad + x_6 * x_7 + x_7 * x_8 + x_8 * x_9 + \ldots \end{aligned} \tag{7}$$

The PR model performed better due to the following reasons [12]

- The best relationship approximation among dependent and independent variables is given by the PR model.
- The PR models can fit functions with broad range.
- Extensive range of curvature can be fit in by PR models.

## 5 Conclusion

As datasets in real world tend to be inconsistent, incomplete & noisy, data pre-processing forms a critical step in knowledge discovery process. It comprises of cleaning, integration, transformation, and reduction of the data. Data cleaning routines can be used to fill in the missing values, smooth noisy data, identify outliers, and correct data inconsistencies. Integration of data combines data from multiple sources that constitutes a comprehensible data store. We have discussed the approach to find missing values of data sets by using technique of Multiple Regressions in which mathematical model helps us to find those corresponding values. The incomplete dataset was constructed using LR, PQR, QR, IR, and PR models. The PR model predicted the missing values accurately than the LR, PQR, QR, and IR models based on the performance metrics $R^2$ and RMSE.

## References

1. Parthasarathy S, Aggarwal CC (2003) On the use of conceptual reconstruction for mining massively incomplete data sets. IEEE Trans Knowled Data Eng 15(6):1512–1521. https://doi.org/10.1109/TKDE.2003.1245289
2. Aggarwal CC (2001) On the effects of dimensionality reduction on high dimensional similarity search. In: Proceedings of the 20th ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, ACM, New York, United States, pp 256–266. https://doi.org/10.1145/375551.383213
3. Han J, Kamber M, Pei J (2011) In: Data mining: concepts and techniques. 3rd edn. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA
4. Sridevi S, Rajaram S, Parthiban C, SibiArasan S, Swadhikar C (2011) Imputation for the analysis of missing values and prediction of time series data. In: International conference on recent trends in information technology (ICRTIT), IEEE, Chennai, India, pp 1158–1163. https://doi.org/10.1109/ICRTIT.2011.5972466
5. Doreswamy IG, Manjunatha BR (2017) Performance evaluation of predictive models for missing data imputation in weather data. In: International conference on advances in computing, communications and informatics (ICACCI), IEEE, Udupi, India, pp 1327–1334. https://doi.org/10.1109/ICACCI.2017.8126025
6. Kotteti CMM, Dong X, Li N, Qian L (2018) Fake news detection enhancement with data imputation. In 16th international conference on dependable, autonomic and secure computing, IEEE, Athens, pp 187–192. https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00042
7. Zeng D, Xie D, Liu R, Li X (2017) Missing value imputation methods for TCM medical data and its effect in the classifier accuracy. In: 19th international conference on e-health networking,

applications and services (Healthcom), IEEE, Dalian, pp 1–4. https://doi.org/10.1109/HealthCom.2017.8210844
8. Zou KH, Tuncali K, Silverman SG (2003) Correlation and simple linear regression. Technical report, Statistical Concepts Series
9. Aditya Shastry K, Sanjay HA, Bhanusree E (2017) Prediction of crop yield using regression techniques. Int J Soft Comput 12:96−102. https://doi.org/10.36478/ijscomp.2017.96.102
10. Wisniewski M (2017) Applied regression analysis: a research tool. J Operational Res Soc 41(8):782–783. https://doi.org/10.1057/jors.1990.106
11. KEEL-dataset repository (2018). https://sci2s.ugr.es/keel/dataset.php?cod=77. Last Accessed 19 June 2018
12. Karama A, Farouk M, Atiya A (2018) A multi linear regression approach for handling missing values with unknown dependent variable (MLRMUD). In: 14th international computer engineering conference (ICENCO), IEEE, Cairo, Egypt, pp 195–201. https://doi.org/10.1109/ICENCO.2018.8636126

# An Improved Stacked Sparse Auto-Encoder Method for Network Intrusion Detection

**B. A. Manjunatha and Prasanta Gogoi**

## 1 Introduction

Today's communication world-wide abundantly depends on the computer network. Generally, to detect the threats from the network used the intrusion detection system (IDS) [1, 2]. The dimensionality reduction (DR) is the most important step in building the IDS system and the most critical part in quantifying the efficiency of the IDS system. DR can transform specific feature vectors into feature vectors which are abstract, which realizes the nonlinear transformation from high dimensional data space to low dimensional data space. The feature or attribute extraction process extracts the reduced and most accurate features in classifying the network data thereby increasing the accuracy, detection rate, and performance of the system. This process extracts the subset by removing the redundant and irrelevant records from the original dataset. More records in the dataset increase the complexity, resource consumption, and difficulty in processing the dataset. This feature extraction process also minimizes the data needed for processing, as the data specific dimensionality reduction will be done, so the system shall be faster and takes fewer resources and time in producing the result.

B. A. Manjunatha (✉) · P. Gogoi
Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore, India
e-mail: manjunatha.ba@nmit.ac.in

P. Gogoi
e-mail: Prasanta.gogoi@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_10

Various supervised and unsupervised machine learning techniques can be used such as principal component analysis (PCA), support vector machine (SVM), random forest (RF) classifiers which have been utilized to detect and classify intrusions. In the proposed model, the KDD-cup 99, NSL-KDD, UNSW-15nb, and NMITIDS datasets considered for building the system will have huge records and features. With the large dataset, and no normalization is processed, then the performance of the IDS will be degraded. The data encoding is the prerequisite to perform the normalization process, the results of normalized data will be used for the feature selection and classification activity.

The paper is prepared as follows: Introduction on intrusion detection and its needs are discussed in the Sect. 1, in Sect. 2, the literature on the intrusion detection are discussed. The methodologies used in our analysis are discussed in the Sect. 3. In Sect. 4, results are discussed. Section 5 conclusion and future work.

## 2 Related Work

In this section, we have discussed most current research works that belong to dimensionality reduction, classification of intrusion detection.

In this paper [3] authors have proposed a model of IDS that targeted more accuracy. The Cuttle Fish feature selection algorithm combined with an extended Chi-square algorithm in this model. During the classification, the existing algorithm namely IAEMSVM is used. The model takes less computation time because of fewer features are used for the analysis. The experimental results of the proposed IDS model are reduced the computation time with fewer features. This proposed IDS model run on KDD Cup 99 data set accuracy is nearly 99%. In this paper [4] authors have shown keen interest in information security. The dataset available for testing an IDS system consists of ineffective and sometimes irrelevant information concerning a particular experiment. The original dataset is a high dimensional database, so perform in testing is very difficult and required high computational time. So, they have removed unwanted information from the original high dimensional database and create a small set of features. The generalized discriminant analysis (GDA) feature reduction method is proposed to overcome the limitations of Principal Component Analysis (PCA). This algorithm increases the accuracy level along with feature reduction. A comparison with Self-Organizing Map (SOM) and C4.5 classifiers shows GDA is outperformed. In this paper [5] authors proposed the rough set theory algorithm used for the reduction of features. The neural network helps in traffic data packet classification. The work carried out on the NSL-KDD dataset. The result shows 80.4% select feature data reductions and detection accuracy is nearly 96.7%. In this paper [6] proposes a novel two-phase deep learning (TSDL) model, given a stacked auto-encoder with a softmax classifier. The model involves two choice stages: an underlying stage liable for arranging system traffic as typical or strange, utilizing a probability score esteem. This is then utilized in an ultimate choice stage as an extra component, for recognizing the typical state and different classes of assaults.

Comparative results show altogether beats existing methodologies, accomplishing high acknowledgment rates 89.134%, for the KDD99 and UNSW-NB15 datasets individually. In this review paper [7] authors discussed feature relevance, evaluation criteria along with various characteristics for feature selection. Various methods are available for feature selection these methods are compared categorically. A legitimate system is examined for the determination of proper feature selection calculations without knowing the data of every calculation. In this paper [8] authors have introduced a new dimensionality reduction technique. The method considered isolated variables, and how this helps in the detection of anomalies for optical emission spectroscopy data. In this paper [9], the authors discussed a novel feature selection methodology. This strategy utilizes include normal of aggregate and each class and afterward applies classifier decision tree calculation for assessing feature reduction technique. This technique is contrasted and three elements choice CFS, IG, and GR. The investigation shows that among exactness and AR esteem is the opposite relationship in our feature selection technique. The most noteworthy precision is 99.794% utilizing 22 features. In this paper [10] a comparative study on various Network IDS methods and deep learning approaches are discussed. The proposed method of deep learning gives better results in terms of detection accuracy and a low false-positive rate. In [11] authors proposed an unsupervised feature reduction. The proposed deep learning method is tested with the KDD Cup99 dataset. The results of the proposed method give better detection and false-negative rate i.e., 97.90% and 2.47%. In [12] authors proposed a deep learning method for designing flexible Network IDS. This method combines a sparse auto-encoder with softmax regression and evaluated with the NSL-KDD dataset. The results show that the standard f-score value is 75.76%. In review paper [13] authors discuss various feature selection techniques are discussed. Novel feature selection techniques are designed due to its importance in network intrusion detection. Various technical aspects are included and categorized these techniques according to their implementation on the KDD-cup99 dataset. In this paper [14] authors proposed a hybrid algorithm based on selected features along with a machine learning approach. This algorithm also addressed the misuse detection, which helps in finding attacks on the NSL-KDD dataset. In this paper [15], the author uses NSL-KDD dataset for designing the IDS model. In this model, five different classifiers with four feature selection techniques are matched on the detection rate. The study compares performances of different combinations of feature selection techniques with classifiers. The feature selection methods namely principal component analysis (PCA), CFS, IGR, and classifiers are Neural Network, K-nearest neighbor, and support vector machine is used. The reduced features acquired in all the blends are IGR feature selection with a good detection rate obtained through the k-NN classifier. In [16] authors proposed a model based on RNN for testing short messages with an automatic security auditing tool. The results show an accuracy rate of 92.7% obtained. In [17] authors designed an Android malware detection framework namely Deep4MalDroid. They have uses three layers of a stacked auto-encoders method for better results. In [18] authors monitor network flow data and design an algorithm that is evaluated on NSL-KDD Dataset. This proposed calculation gives a precision of 75.75% utilizing six fundamental features.

The literature reviews show various classification and feature selections methodologies are uses for intrusion detection in the intruder detection system by using the KDD-cup 99 dataset and its advanced version of dataset NSL-KDD. Many works are presented based on numerous data mining methods and deep learning methodologies. Although the accuracy of the detection level of anomalies is good still there is always a scope of further improvement that is possible with respect to intrusion detection accuracy and other parameters.

## 3 Proposed Framework

The proposed method in this research work uses a deep learning technique for the IDS system. The proposed method uses existing deep learning techniques [19] used to achieve dimensionality reduction in the intrusion dataset. Sparsity constraints are added to the sparse auto-encoder method to develop improved dimensionality reduction. Hence, the sparse penalty introduced by sparse coding in the hidden layers of auto-encoder. This promotes the auto-encoder to take up and produce brief and effective low-dimensionality information features and quickens the classifier process on the intrusion dataset. So, the proposed framework of improved stacked sparse auto-encoder (ISSAE) is superior to the known techniques.

A well-known method as of now used unsupervised inside deep learning research is auto-encoder [1]. This method consists of three sections: encoder, decoder, and code. This method uses encoder and decoder are a particular sort of feed-forward neural systems where the info is equivalent to the yield. The architecture of the auto-encoder shown in Fig. 1 [19]. The code is a solitary layer of an artificial neural network (ANN) with our preferred dimensionality. The quantity of nodes in the code layer is a hyper-parameter that we set before preparing the auto-encoder.

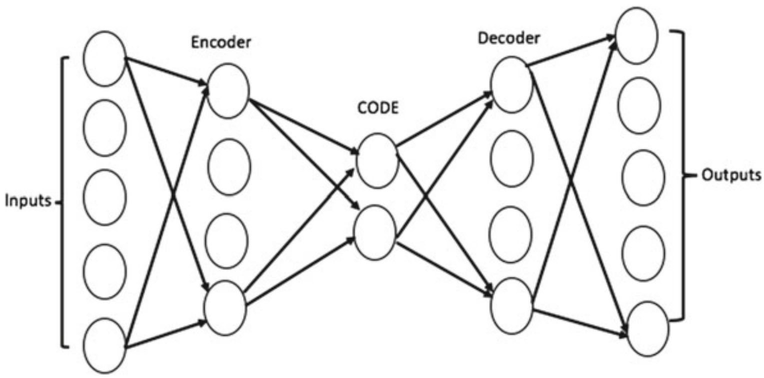



**Fig. 1** Architecture of Auto-encoder

### 3.1 Sparse Auto-Encoder for Dimensionality Reduction:

The input data gets as simple as copying in an auto-encoder to the output layer and this a problem as it doesn't extract any meaningful information. This promotes the auto-encoder to take up and produce concise and reduce the dimensionality of features. In other words, the sparse constraints are applied to improved precise input features. It considered the average activation function, say $\widehat{\rho_j}$ which is given in the hidden layer shown in Eq. (1).

$$\widehat{\rho_j} = \frac{1}{m}\sum_{i=1}^{m}\left[a_j^{(2)}(x^{(i)})\right] \tag{1}$$

The auto change which might happen and which the hope is that the average activation function $\widehat{\rho_j}$ approaches $\rho$ which is close to zero.

To achieve this purpose, the Kullback–Leibler (KL) divergence is added as a regularization term to the squared error function of the auto-encoder and this is given in Eq. (2).

$$\mathrm{KL}(\rho||\widehat{\rho_j}) = \rho\log\frac{\rho}{\widehat{\rho_j}} + (1-\rho)\log\frac{1-\rho}{1-\rho_j} \tag{2}$$

This is the combination of entropy and as well as cross-entropy. KL divergence transforms similarity in data points to joint probabilities. The addition of this term to error function and how it benefits dimensionality reduction is shown in later sections.

### 3.2 Stacked Sparse Auto-Encoder

Stacked sparse auto-encoder (SSAE) is composed of numerous sparse auto-encoders. The yield of the past layer is given as the contribution of the following layer of the self encoder and this is the way the feed-forward network happens. It makes the auto-encoder realize more accurate features and minimizes redundancy. The weights and biases of the network are being reduced and a minimal squared error function's value is obtained. This means optimizing the weights and bias will yield good results. The greedy layer-wise pre-preparing strategy is utilized to successively prepare each layer of ISSAE and Adam optimization algorithm is used to implement dynamic adjustment of different parameters.

### 3.3 Proposed Framework of Improved Stacked Sparse Auto-Encoder Dimensionality Reduction Method

Initially, the feature extraction starts when the pre-processed dataset is fetched and reduced all the symbols to numerical form. The dataset KDD-cup 99, NSL-KDD has 41 classification features, UNSW-15nb has 49 classification features, and NMITIDS dataset has 31 classification features. The proposed network intrusion detection framework is shown in Fig. 2 this explains the effective intrusion detection using ISSAE dimensionality reduction method. Our number of inputs and create a blueprint of the auto-encoder neural network by making a list of units in input, hidden and output units. Initializing learning rate, cost, number of epochs, the batch size for the auto-encoder are done. The encoder and decoder parameters are initialized too. We used truncated normal distribution and random normal distribution for declaring weights and bias vectors. Using the blueprint of the layers initialized we calculate the activation functions of each neuron or unit in each layer and then calculate the sigmoid function for each iteration. Now the predicted value and the true value are being compared using squared error function which is our cost function. After the encoding stage sparsity constraints are being applied by calculating the average activation function and finding out the Kullback–Leibler divergence. For each epoch in this proposed work cost function is directly proportional, the main intent is to minimize the cost function. As discussed earlier the sparsity parameter and the process of stacking when come together will make our squared error function as minimal as possible. The Kullback–Leibler (KL) divergence is additional as a regularization term to the squared error function of the auto-encoder to accomplish the sparsity.

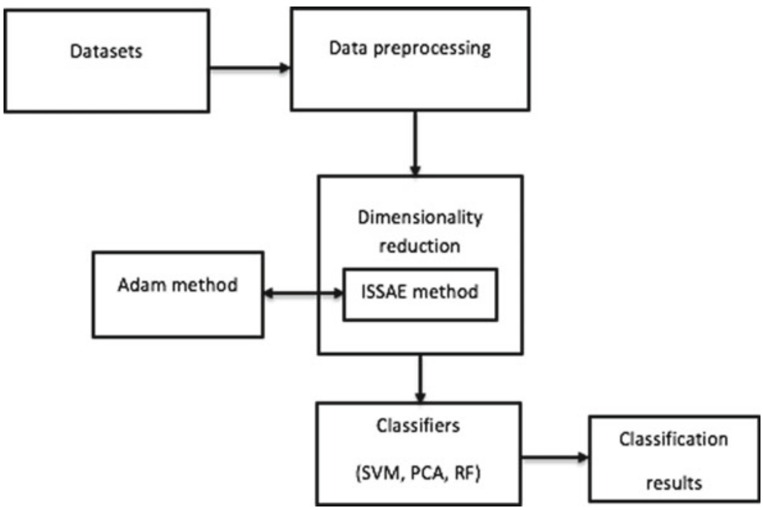



**Fig. 2** Network intrusion detection framework

The strength of the sparse factor control by using weight factor error ($\mu$). The error function show in Eq. (3) then becomes:

$$J_{\text{Sparse}}(w, b) = J(w, b) + \mu \sum_{\text{j=1}}^{\text{m}} \text{KL}(\rho || \widehat{\rho_j}) \tag{3}$$

where, $\mu$ is a weighting factor,

$m$ is hidden units.

To forestall overfitting the weight constriction things is added for error function,$\lambda$ is the coefficient of the weight shown in Eq. (4).

$$J_{\text{Sparse}}(w, b) = J(w, b) + \boldsymbol{\mu} \sum_{j=1}^{m} \text{KL}(\rho || \widehat{\rho_j}) + \frac{\lambda}{2} \sum_{r=1}^{3} \sum_{i=1}^{m} \sum_{j=1}^{m+1} (w_{ij}^r) \tag{4}$$

Secondly, the greedy layer-wise pre-preparing technique is utilized to prepare each layer of SSAE to gain optimized weights and bias estimations of the whole stacked sparse network. At that point the error back propagation technique is utilized to calibrate until the result of the error function between the input data and the output data satisfies the expected requirements. Therefore, the update process of weight and bias is carried on Eqs. (5) and (6).

$$(\boldsymbol{w}_{ij}^{K}) = (\boldsymbol{w}_{ij}^{K}) - \boldsymbol{\eta} \frac{\boldsymbol{\theta}}{\theta \boldsymbol{w}_{ij}^{k}} \boldsymbol{J}(\boldsymbol{w}, \boldsymbol{b}) \tag{5}$$

$$\left(b^r\right) = \left(b^r\right) - \boldsymbol{\eta} \frac{\boldsymbol{\theta}}{\theta b^r} \boldsymbol{J}(\boldsymbol{w}, \boldsymbol{b}) \tag{6}$$

The following is the ISSAE based feature extraction algorithm, which will be used in producing the low dimensional dataset.

***Algorithm 1: ISSAE: Improved Stacked Sparse Auto-encoder***

**Input:** ninput, number of features; input_data, dataset; learning_rate, learning rate of autoencoder training_epochs = 10, number of times the dataset is provided batch_size, size of data in a batch sent
nlayers, list - units in each layer
y_true, memory allocation of size ninput
roe = 0.05, value
**Output:** cost, value of error function for each epoch
**Begin:** for range from 1 to epoch
Taking input_data of batch_size as array of 2D
for x in range from 1 to len(nlayers)
Declare encoder = [
'W': truncated normal distribution ([nlayers[x-1],nlayers[x]],)
'b': random normal distribution ([nlayers[x]])
];
end
for x in range from len(nlayers) to 1
Declare decoder = [
'W': truncated normal distribution ([nlayers[x],nlayers[x-1]],)
'b': random normal distribution ([nlayers[x-1]])
];
end
//Computing encoding operation
for i in range of len(encoder):
encoder_op = sigmoid(sum(matmul(layers[i],encoder[i]['W']),encoder[i]['b']));
end
//Computing KL Divergence
if(j > = 2)
KL = $\left(roe * log \frac{roe}{roe_{cap}}\right) + (1 - roe)(\frac{1-roe}{1-roe_{cap}})$;
Cost = cost + KL;
endif
//Computing decoding operation
decoder_op = [encoder_op] + []
for i in range of len(decoder):
decoder_op = sigmoid(sum(matmul(layers[i],encoder[i]['W']),encoder[i]['b']));
end
y_pred = decoder_op;
cost = (y_true-y_pred)$^2$/(2*ninput);
//Apply Adam Optimization Algorithm
optimizer = AdamAlgorithm(learning_rate,cost);
//next section
print epoch and cost // show in Fig. 3
end
end

**Algorithm 2: Adam optimization algorithm**

**Require:** $\alpha$: Stepsize
**Require:** $\beta_1, \beta_2 \in [0, 1)$: Exponential decay rates for the moment estimates
**Require:** $f(\theta)$: Stochastic objective function with parameters $\theta$
**Require:** $\theta_0$: Initial parameter vector
  $m_0 \leftarrow 0$ (Initialize 1st moment vector)
  $v_0 \leftarrow 0$ (Initialize 2nd moment vector)
  $t \leftarrow 0$ (Initialize timestep)
  **while** $\theta_t$ not converged **do**
    $t \leftarrow t + 1$
    $g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$ (Get gradients w.r.t. stochastic objective at timestep $t$)
    $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first moment estimate)
    $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)
    $\widehat{m}_t \leftarrow m_t/(1 - \beta_1^t)$ (Compute bias-corrected first moment estimate)
    $\widehat{v}_t \leftarrow v_t/(1 - \beta_2^t)$ (Compute bias-corrected second raw moment estimate)
    $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m}_t/(\sqrt{\widehat{v}_t} + \epsilon)$ (Update parameters)
  **end while**
  **return** $\theta_t$ (Resulting parameters)

Now we have sparse requirements in the ISSAE arrange, we need to utilize diverse learning rates for various parameters, such as diminishing the recurrence of updates for parameters. Classic gradient descent calculations incorporate mini-batch usual gradient descent, in which it is hard to pick a reasonable learning rate since all the parameters are to be updated and get a local minimum. For this purpose, to prepare a superior SSAE to organize, the versatile second estimation gradient descent method calculation is utilized to accomplish the dynamic adaptive adjustment of various parameters. This calculation actualizes dynamic adjustment of various parameters by calculating the gradient first-order request estimate mt and second-order request estimate vt parameters in Algorithm 2. The effect on cost function due to these dynamic adjustments are effective.

## 3.4 *Classification*

After dimensionality of the dataset is reduced from the feature extraction process, the support vector machine (SVM) [20], random forest (RF) [21, 22], and principal component analysis (PCA) [23] techniques are used to classify the network intrusion data records.

### 3.4.1 SVM Classifier

Support Vector Machine (SVM): This technique is very useful in classification, regression, and outlier detection. This can be used as a linear and non-linear classifier. This model can be easily used if any dataset having features and class labels.

In a linear classifier model, training examples are plotted in space. This technique predicts a straight hyper-plane diving two classes. The objective is to construct a hyperplane that characterizes all train vectors in two classes. This hyper-plane is known as the maximum margin hyper-plane.

But in Non-linear SVM used in a dataset which is generally dispersed up to some level. In such a data set kernel trick is used to maximum margin hyperplanes and information focuses are plotted in higher dimensional space as compared to linear SVM.

### 3.4.2 RF Classifier

Random forest (RF) is utilized for categorization and regression analysis. RF is an assortment of numerous decision trees in the preparation stage and yield class names those have the majority vote [24]. RF accomplishes high classification precision and can deal with anomalies in the information dataset. RF is utilized in this work since it is less susceptible to over-fitting and it has recently indicated great categorization results. A pre-handled example of n tests is taken care of to the arbitrary woodland classifier. RF makes n various trees by utilizing various element subsets. Each tree delivers a characterization result, and the aftereffect of the order model relies upon the greater part casting a vote. The example is allocated to the class that acquires the most raised vote scores. The recently achieved classification results demonstrate that RF is sensibly appropriate in the intrusion dataset.

### 3.4.3 PCA Classifier

PCA is used for feature selection, not as a classifier but if it is assumed that the same variable used to fit place new observations the PCA is measured on new points. Then one can just keep the new points at the weighted sum of the variable scores, weights given by the data. It can also be used with a good classifier and when visualized it can be less appropriate to draw a line and classify.

Initially, the data are thoroughly tested after that give training to the data set. Once training of dataset is over a deep-learning and auto-encoder is used to remove the redundancy from the dataset. After removing the redundancy, we have applied the following classification algorithms Support Vector Machine (SVM), Random forest, and principal component analysis (PCA) to find the accuracy of intrusion detection.

## 4 Results

To evaluate effectively of classification methods, we have computed precision, recall, F-score, false-positive rate, and misclassification/error rate. On the off chance that improving the precision diminishes recall and the other way around. Apart from

**Fig. 3** Confusion matrix

| | **Predicted No** | **Predicted Yes** |
|---|---|---|
| **Actual No** | TN | FP |
| **Actual Yes** | FN | TP |

these, we have also minimized the error rate and false positive rate from our proposed methods.

A benchmark dataset for network intrusion detection is KDD-cup 99 [25], NSL-KDD [26] and UNSW NB 15 [27] datasets are taken to evaluate the performance of our proposed IDS. The NMITIDS dataset has been prepared in the NMIT research lab using a testbed setup consisting of high-end servers, cisco router, cisco switch, and the number of client systems. Set up hardware and software environments to create our dataset in the Information science research lab in the year 2019 i.e., NMITIDS dataset [26]. It consists of 31 features, five attack types such as Ncrack, Medusa, Hydra, Mydoom and DBotnet. It has *897,182* lakh records, it has both training and testing sub-datasets and standard service protocols such as TCP, UDP, ICMP and SSH etc., available in NMITIDS dataset.

## *4.1 Confusion Matrix*

A confusion matrix is a standard measure for the results of the classifiers. It is a table that is regularly utilized to describe the display of classifier results of test information for that the genuine values are known as shown in Fig. 3. There are two potential anticipated classes: yes and no.

True Negative Rate (TNR)—The TNR counter is incremented one when the dataset record actual class is abnormal and will also be classified as abnormal.

True Positive Rate (TPR)—The counter is incremented if an actual and classified class is the same (normal) for the dataset.

False Positive Rate (FPR)—The FPR counter is incremented if an actual abnormal class record is classified as a normal record.

False Negative Rate (FNR)—The FNR counter is incremented when a normal class record is classified as an abnormal record.

**Precision:** The extent of positive IDs is really right. Precision is characterized as follows:

$$\textbf{Precision} = \frac{TP}{(TP + FP)} \tag{7}$$

**Recall:** The extent of real positives is recognized accurately. A recall is characterized as follows:

**Table 1** Comparison of results for all classification algorithms on KDD-cup 99 dataset

| Classification methods | Parameters | | | | |
|---|---|---|---|---|---|
| | Precision | Recall | F-score | False positive rate | Misclassification or error rate |
| PCA | 0.9201 | 0.9450 | 0.9323 | 0.0142 | 0.0798 |
| RF | 0.9586 | 0.9514 | 0.9549 | 0.0036 | 0.0060 |
| SVM | 0.9247 | 0.9589 | 0.9360 | 0.0045 | 0.0161 |

$$\textbf{Recall} = \frac{TP}{(TP + FN)} \tag{8}$$

**F-Score:** This is a weighted normal of the true positive rate (recall) and precision. The F-score calculates the harmonic mean of precision and recall.

$$\boldsymbol{F - score} = 2\left(\frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}\right) \tag{9}$$

$$\textbf{False Positive Rate} = \frac{\text{FP}}{\text{Actual No}} \tag{10}$$

$$\textbf{Misclassification or error rate} = \frac{(FP + FN)}{\text{Total}} \tag{11}$$

Table 1 shows the overall results obtained in the form of Precision, Recall, F-score, False Positive Rate, and Misclassification or error rate through our proposed framework. In Table 1, the highest precision value achieved in RF is 95.86%. The recall value is quite the same in RF and SVM nearly 95.14 and 95.89%. The value of F-Score is highest in RF i.e. nearly 95%. The value of the false positive rate is better in RF and PCA as compared to SVM. The misclassification or error rate in PCA is high i.e. nearly 8%. This KDD-cup 99 dataset uses less cost function values and minimum epochs train shown in Fig. 3a.

From the Table 2, shows the overall results obtained in the form of Precision, Recall, F-score, False Positive Rate, and Misclassification or error rate through our proposed framework. In Table 2, the highest precision value achieved in RF

**Table 2** Comparison of results for all classification algorithms on NSL-KDD dataset

| Classification methods | Parameters | | | | |
|---|---|---|---|---|---|
| | Precision | Recall | F-score | False positive rate | Misclassification or error rate |
| PCA | 0.9348 | 0.9048 | 0.9243 | 0.0452 | 0.0998 |
| RF | 0.9647 | 0.9865 | 0.9753 | 0.0089 | 0.0667 |
| SVM | 0.9034 | 0.8963 | 0.8998 | 0.1001 | 0.1061 |

is 96.47%. The recall value achieved low in SVM and PCA nearly 89.63 and 90.48% (Fig. 4).

The value of F-Score is highest in RF i.e. nearly 98%. The value of the false positive rate is better in RF as compared to SVM and PCA. The misclassification or error rate in SVM is high i.e. nearly 10%. This NSL-KDD dataset uses less cost function value and minimum epochs train shown in Fig. 3b.

Table 3 shows the overall results obtained in the form of Precision, Recall, F-score, False Positive Rate, and Misclassification or error rate through our proposed framework. In Table 3, the highest precision value achieved in RF is 95.96%. The recall value is quite the same in PCA and SVM nearly 90.70 and 90.23%. The value of F-Score is highest in RF i.e. nearly 95%. The value of the false-positive rate is better in RF as compared to SVM and PCA. The misclassification or error rate in

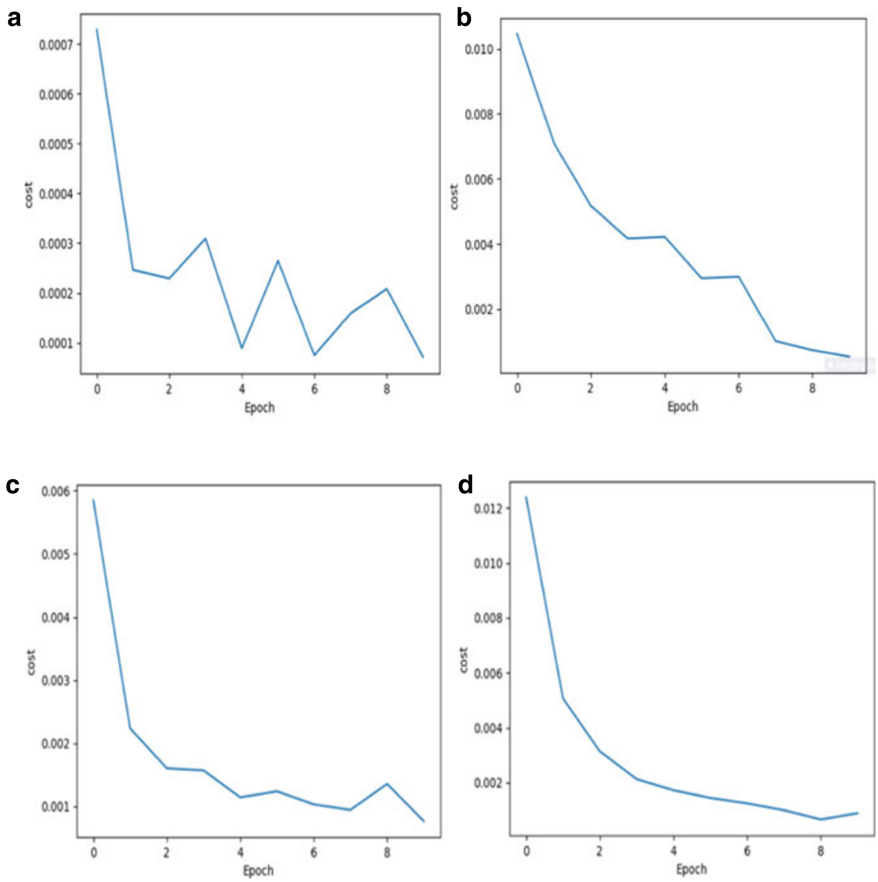



**Fig. 4** **a** Cost versus epoch graph on KDD-cup 99, **b** cost versus epoch graph on NSL-KDD, **c** cost versus epoch graph on UNSW 15nb, **d** cost versuss epoch graph on NMITIDS

**Table 3** Comparison of results for all classification algorithms on UNSW-15nb dataset

| Classification methods | Parameters | | | | |
|---|---|---|---|---|---|
| | Precision | Recall | F-score | False positive rate | Misclassification or error rate |
| PCA | 0.9488 | 0.9070 | 0.9277 | 0.0892 | 0.0799 |
| RF | 0.9596 | 0.9426 | 0.9511 | 0.0123 | 0.0670 |
| SVM | 0.8978 | 0.9023 | 0.8999 | 0.2754 | 0.3521 |

**Table 4** Comparison of results for all classification algorithms on NMITIDS dataset

| Classification methods | Parameters | | | | |
|---|---|---|---|---|---|
| | Precision | Recall | F-score | False positive rate | Misclassification or error rate |
| PCA | 0.9123 | 0.9098 | 0.9110 | 0.0623 | 0.0543 |
| RF | 0.9612 | 0.9520 | 0.9567 | 0.0891 | 0.0673 |
| SVM | 0.9389 | 0.9265 | 0.9325 | 0.0925 | 0.0943 |

SVM is high i.e. nearly 35%. This UNSW 15nb dataset uses less cost function values and minimum epochs train shown in Fig. 3c.

From the Table 4, shows the overall results obtained in the form of Precision, Recall, F-score, False Positive Rate, and Misclassification or error rate through our proposed framework. In Table 4, the highest precision value achieved in RF is 96.12%. the recall value is low in PCA nearly 0.9098. The value of F-Score is highest in RF. The value of the false positive rate is better in PCA as compared to SVM and RF. The misclassification or error rate is high in SVM i.e. nearly 5%. This NMITIDS dataset uses less cost function values and minimum epochs train shown in Fig. 3d.

## 5 Conclusion

The proficiency and effectiveness of NIDS be contingent on robust feature extraction and classification algorithms. Offering the Auto-encoder model by conjoining stacking and sparsing constraints in extracting the best optimal feature set. The Stacked Sparse Auto-encoder is an improved feature extraction model from the present systems. As discussed in Sects. 3 and 4, the projected system was built and evaluated for KDD-cup 99, NSL-KDD, UNSW-15nb, and NMITIDS datasets records. The improved stacked sparse auto-encoder approach provides the best finest feature set with a reduced number of records among 41 features for the KDD-Cup-99 and NSL-KDD datasets respectively. Similarly, 49 features for the UNSW-15nb dataset and 31 features for the NMITIDS dataset too. The classification performance facts also display minimum cost function and fewer epoch trains are in-line with the standard model. Proposed system results are paralleled with the other state of the art methods and results display that the proposed model got better enhancements over the existing orientation model.

The proposed system is supple enough to house any new dataset, so with minimal modifications, the new dataset shall be evaluated. The RF, SVM, and PCA classification model was used in this work for attack classification. Other classification models also can be strained for improved performance results. Also in factual time situations, the data is more of a non-linear type, so the system shall be boosted for nonlinear or binding IDS approaches for improved feature and classification performance.

## References


1. Shone N, Ngoc TN, Phai VD, Shi Q (2018) A deep learning approach to network intrusion detection. IEEE Trans Emerg Topics in Comput Intell 2(1):41−50
2. Manjunatha BA, Gogoi P (2016) Anomaly based intrusion detection in mixed attribute dataset using data mining methods. J Artif Intel 9(1–3):1–11
3. Rajesh Kambattan K, Manimegalai R, Ganapathy S (2017) An increment feature selection approach for intrusion detection system in MANET. Int J Res Appl Sci Eng Technol (IJRASET) 5(I):325–329
4. Singh S, Silakari S, Patel R (2011) An efficient feature reduction technique for intrusion detection system. In: International conference on machine learning and computing, IPCSIT vol 3. IACSIT Press, Singapore, pp 147–153
5. Sadek RA, Sami Soliman M, Elsayed HS (2013) Effective intrusion intrusion detection system based on neural network with indicator variable and rough set reduction. IJCSI Int J Comput Sci 10(6) (2):227–233
6. Sisiaridis D, Markowitch O (2017) Feature extraction and feature selection: reducing data complexity with apache spark. Int J Netw Secur Its Appl (IJNSA) 9(6):39–51
7. Kumar V, Minz S (2014) Feature selection: a literature review. Smart Comput Rev 4(3):211–229
8. Puggini L, Macloone S (2016) Feature selection for intrusion detection using optical emission spectroscopy. IFAC-Papers On Line 49–5:132–137
9. Chae H-S, Choi SH (2014) Feature selection for efficient intrusion detection using attribute ratio. Int J Comput Commun 8
10. Dong B, Wang X (2016) Comparison deep learning method to traditional methods using for network intrusion detection. In: Proceedings 8th IEEE international conference communications software networking, Beijing, China, June 2016, pp 581–585
11. Alrawashdeh K, Purdy C (2016) Toward an online intrusion intrusion detection system based on deep learning. In: Proceedings 15th IEEE international conference machine learning applications, Anaheim, CA, USA, December 2016, pp 195–200
12. Javaid A, Niyaz Q, Sun W, Alam M (2016) A deep learning approach for network intrusion detection system. In: Proceedings 9th EAI international conference bio-inspired information communications technology, pp 21–26. Online. Availablehttps://doi.org/10.4108/eai.3-12-2015.2262516
13. Rais HBM, Mehmood T (2016) Feature selection in intrusion detection, state of the art: a review. J Theoret Appl Info Technol 94(1):30–43
14. Sasan HPS, Sharma M (2016) Intrusion detection using feature selection and machine learning algorithm with misuse detection. Int J Comput Sci Info Technol (IJCSIT) 8(1)
15. Biswas SK (2018) Intrusion detection using machine learning: a comparison study. Int J Pure Appl Math 118(19):101–114

16. You L, Li Y, Wang Y, Zhang J, Yang Y (2016) A deep learningbased RNNs model for automatic security audit of short messages. In: Proceedings 16th international symposium communications information technology, Qingdao, China, September 2016, pp 225–229
17. Hou S, Saas A, Chen L, Ye Y, Deep4MalDroid: a deep learning framework for android malware detection based on linux kernel system call graphs. In: Proceedings IEEE/WIC/ACM international conference web intelligence workshops, Omaha, NE, USA, October 2016, pp 104–111
18. Tang TA, Mhamdi L, McLernon D, Zaidi SAR, Ghogho M (2016) Deep learning approach for network intrusion detection in software defined networking. In: Proceedings international conference wireless networks mobile communication, October 2016, pp 258–263
19. https://towardsdatascience.com/applied-deep-learning-part-3-autoencoders-1c083af4d798 Last Accessed on September 2018
20. Support Vector Machine (2018). http://dataaspirant.com/2017/01/13/support-vector-machine-algorithm/ Last Accessed on November 2018
21. https://www.datascience.com/resources/notebooks/random-forest-intro. Last Accessed on November 2018
22. Chen X, Ishwaran H (2012) Random forests for genomic data analysis. Genomics 99:323–329
23. Principal Component Analysis (2019). ftp://statgen.ncsu.edu/pub/thorne/molevoclass/AtchleyOct19.pdf
24. Garcia Cordero C, Hauke S, Muhlhauser M, Fischer M (2016) Analyzing flow-based intrusion intrusion detection using replicator neural networks. In: Proceedings 14th annual conference privacy, security trust, Auckland, New Zeland, December 2016, pp 317–324
25. https://kdd.ics.uci.edu/databases/kddcup99/task.html
26. https://www.nmit.ac.in/department/information-science-research.html
27. Moustafa N, Slay J (2015) UNSW-NB15: a comprehensive data set for network intrusion detection systems. In: Military communications and information systems conference (MilCIS), IEEE, pp 1–6

# A Node Quality Based Cluster Header Selection Algorithm for Improving Security in MANET

**S. Muruganandam and J. Arokia Renjit**

## 1 Introduction

MANET stands for Mobile Ad-hoc Networks, also known as Wireless Ad-hoc Networks, it does not requires a fixed infrastructure support for transforming data packets between two nodes. MANET becomes an attractive technology due to increasing demands of fast and quality data sharing, easy to deployment [1]. The important attribute of MANET is shared and self organized in order to perform the defined network functionalities through the existing sensor nodes and hence the collaboration among nodes is an essential factor for providing a stable communication more efficiently. The major challenge in the design of MANET includes security, energy consumption, Node Clustering methods and Cluster Node (CH) selection [2]. All these challenges can be resolved by developing efficient clustering algorithm for selecting a sustainable node as a cluster head node. The Cluster Head (CH) node is authorized node for monitoring the entire node behavior as well as isolating the misbehavior nodes and also all security related functions are initiated by CH node. Figure 1 shows that the architectural model of Mobile Ad-hoc network, every node in the network is grouped to form a cluster, each cluster having the cluster head node connected with wireless link. The base station manages all cluster networks with its signal range.

**Various Clustering Methods in MANET** The clustering methods are categorized based on the approach used for clustering [3].

(a) *Heuristic Schemes*

S. Muruganandam (✉)
Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

J. Arokia Renjit
Department of Computer Science and Engineering, Jeppiaar Engineering College, Chennai, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_11

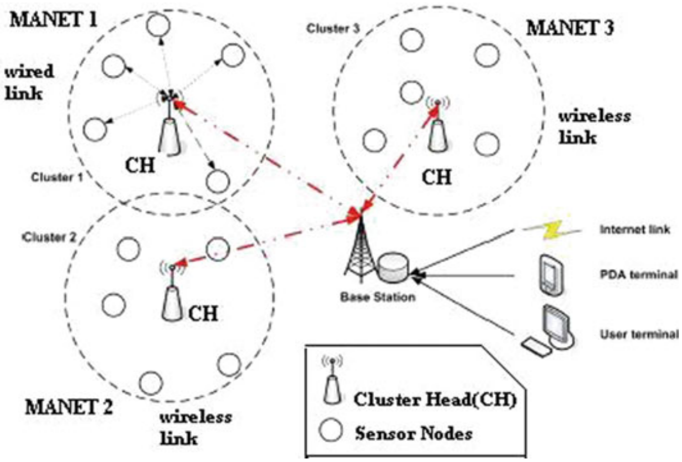



**Fig. 1** MANET architecture model

This kind of clustering algorithm does not based on measurements of mobile nodes; it gives the optimal solution at a reasonable run time. Eg: Linked Cluster Algorithm (LCA), highest connectivity of a node, Max–Min D-Cluster algorithm.

(b) *Weighted Schemes*

This algorithm is based on the node quality, several parameters are used to assess the efficiency of a node, based on this node value, clustering can be performed. Eg: Weighted Clustering Algorithm (WCA).

(c) *Hierarchical Schemes*

In this algorithm a node clustering can be performed based on the protocol used in the MANET. Eg: LEACH, TL-LEACH, EECS, and HEED.

(d) *Grid Schemes*

This algorithm performs a node clustering bases on the energy levels of nodes. Eg: PEGASIS, GROUP.

This paper proposes node efficiency based Cluster Head (CH) selection algorithm based on K-means clustering algorithm and weighted clustering Algorithm (WCA) [4, 5]. This algorithm effectively selecting a high sustainable cell rate (SC) node a Cluster Head (CH). The sustainable cell rate (SC) of a node can be calculated based on the important attribute of a node such as energy consumption, mobility of a node, distance between the nodes, residual energy and link quality of a node, using these parameter a sustainable cell rate (SC) of all nodes will be computed individually. After computing SC rate of all nodes, the priority list of possible Cluster Head (CH) nodes are created by the network admin or service provider. From this approach a node having high Sustainable Cell (SC) rate as a Cluster Header [5]. If the present

CH node reaches the end state it will call and assign a next priority SC rate node as Cluster Head (CH), so the time taken for reassigning a node will be reduced. This proposed algorithm will increase the network lifetime and improve performance of network at considerable level while comparing the existing cluster head selection algorithms.

*MAC layer*: The MAC Layer is important for forwarding the data packets to and from one hub to another hub in dynamic mobile Ad-hoc networks. Before the data transmission starts the network admin or service provider must ensure the fair sharing of network access and resolving the design issues of MAC layer. Two main design issues of MAC layers are hidden terminals and exposed terminals implementing a CSMA/CD protocol to avoid collusion in a common communication medium, designing an efficient MAC layer for improving the network stability. Eliminating the problem of MAC layer is a first factor while performing a cluster head selection process.

*Network model specification*: The Mobile Ad-hoc networks are represented as an undirected graph.

Let it be, $M(G) = (V, E)$ Created by nodes and communication link [6, 7]. Now a set of nodes $V_i$ and a set of communication link $E_i$ are commonly expressed as $V$ and $E$. $|E_i|$ represents the modification due to creations and deletions of the links although there is no impact on cardinality of $V$ ($|V|$). Degree of node $V_i$ is computed as:

$$\mathrm{Deg}(V_i) = \left\{V_j, \text{such that } \mathrm{Dist}\left(V_i, V_j\right) < r_{vi}\right\}$$

where $r_{vi}$ is the communication range and $\mathrm{Dist}(V_i, V_j)$ is the mean distance computed from $V_i$ to $V_j$.

The clustering algorithm must satisfy the following three conditions

(a) Two CH cannot be a nearest nodes
(b) Every node contains minimum one CH node as its neighbor
(c) The weightage value of a CH node is must be greater than all other nodes in the network.

The Intrusion Detection System (IDS) is used for identifying different types of network attacks. The researchers proposed a IDS for monitoring a network attacks using an encrypted data and SVM [8, 9]. There is a need for developing a method for trust based clustering solutions in MANET for improving a network security [10].

## 2 Related Work

To achieve energy efficiency and security in MANET, the energy aware suspicious node identification for preventing packet transmission misconduct attacks [11].

Trusted agent based low process surprise validation for malicious node identification is developed. This method identifies the malicious node by packet transmission rate, node angle method, node response time and secured data transmissions. The cluster Header node is elected based on the remaining energy and degree of connectivity of a node. The trusted mobile agent improves the security of network by observing the packet transmission [12]. Trust based clustering method is implemented for providing secure communication in MANET. The multi value fuzzy logic is used for classification of mobile nodes. The trust value of a node is computed by nodes behavior [13]. The contexts aware routing protocol is implemented for making routing decisions based on various network quality attributes. This proposed method is developed using K- means clustering algorithm with unsupervised machine learning algorithm [14]. The node quality based clustering algorithm is developed for improving the energy efficiency and stability of the network [15]. Trust management scheme in wireless Ad-hoc network is implemented by combining two levels of security, computing reliability of a node and node certificate authentication method [16]. The new algorithm is developed to reduce black hole and gray hole attack. This method reduces the packet drop rate and increasing network throughput. The proposed algorithm improves network security [17]. The support vector based algorithm is for analyzing node features and classifying the node as trusted node and malicious node. This algorithm improves the trust evaluation performance [18]. A new reputation based routing protocol is developed for enhancing the performance of MANET [19]. The trusted node identification algorithm for improving energy efficiency of MANET. In this approach the random nodes are grouped to form a cluster models and trusted nodes are identified from each cluster by using detailed analytical algorithm such as improved optimization algorithm [20]. The node reputation based method to avoid suspicious nodes in wireless sensor networks. This approach uses the past history, while packet transmission. The node reputation is tested to select the best reliable route. This algorithm effectively removes the effects of suspicious nodes and enhances the network security [21]. The trust value updating algorithm for multipath routing algorithm for cluster based wireless sensor network. This algorithm selects a node having highest trust value as a Cluster Header (CH).The CH node is modified dynamically based on updated trust value [22].

## 3 Proposed Architecture

The proposed architecture model in Fig. 2 shows the working principal of an implemented algorithm, selecting a Cluster Head (CH) node in a mobile network is performed by using various clustering criteria. All nodes participated in a mobile network, must compute the value of four attributes such as energy consumption, mobility of a node, Distance between nodes, Delay in transmission; individually by using separate functions, adding all these value of every mobile node for getting

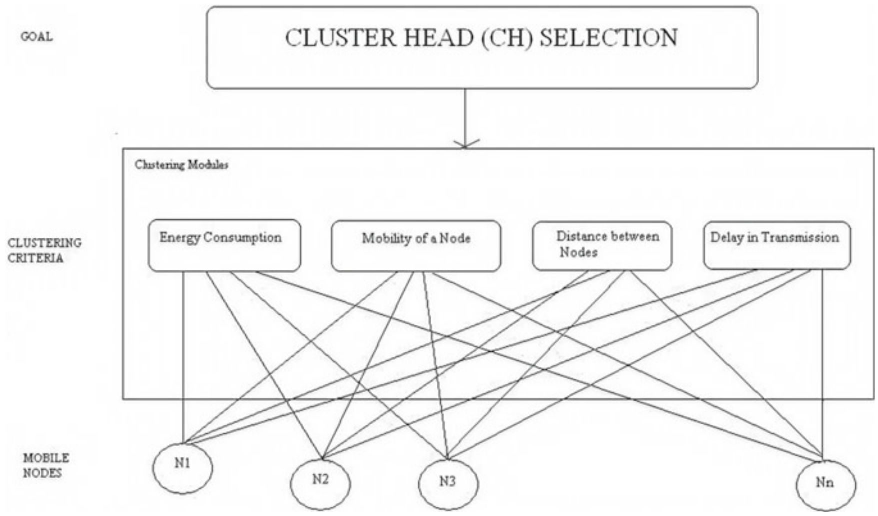



**Fig. 2** Proposed architecture model

cumulative value of a node, based on these value factor, a node having high Sustainable Cell (SC) rate is selected as a cluster header. This proposed architecture is explained by following proposed methods.

## 4 Proposed Methodology

The proposed methods aim to improve the stability of a network by using combined weight metrics model.

The weight of each node is calculated by using the equation

$$W_v = W_1.\,E + W_2.M + W_3.D + W_4.TD \tag{1}$$

$W_1$.E—Energy consumption
$W_2$.M—Mobility of a node
$W_3$.D—Average distance between the CH and base stations
$W_4$.TD—Transmission Delay.

The proposed method consists of four modules for calculating the values of each node.

(a) Energy Consumption
(b) Mobility of a node
(c) Distance between CH and base stations
(d) Delay in Transmission.

## 4.1 Energy Consumption of a Node

A energy consumptions of a mobile node during the data transfer is depends upon several factors such as the distances between the sender node and receiving node, the number of hubs between the origin node and target node, volume of data transferred and the no of bits delivered per second.

### 4.1.1 Average Energy Consumptions

It is calculated by the proportion of total amount of energy utilized by each node in the network is divided by its initial energy.

$$\text{Average energy consumptions} = \frac{\Sigma\ \text{Energy utilized}}{\Sigma\ \text{Initial energy}} * 100 \tag{2}$$

Energy consumption of a node also depends upon the distance of data transmission, so the path length must be computed.

**Average Path length** The path length between two vertices $u$ and $v$ in a graph $G$.

d($u$, $v$) denotes the path length between $u$ and $v$.

The average path length of a connected graph between two vertices $u$ and $v$ in a graph $G$.

$$L(G) = \frac{1}{n(n-1)} \Sigma \mathrm{d}(u, v) \tag{3}$$

## 4.2 Mobility of a Node

The Mobility of a node is defined as change of one location to another location. The mobility of a Node, $A$ at given time $T$ in the mobile ad hoc network is represented as changes of its neighboring nodes correlated to the previous state at time $T - nT$. Thus, nodes that join or/and leave the neighboring of the node $A$ will impact the evaluation of its mobility of a node $A$ at time $T$, by the following formula

$$M_A(t) = \alpha \frac{\text{No of Nodes Out } (T)}{\text{No of Nodes } (T - nT)} (1 - \alpha) \frac{\text{No of Nodes In}(T)}{\text{No of Nodes } (T)} \tag{4}$$

where,

No of Nodes In($T$): The sum of nodes that connected the bounds of $A$ during at interval $[T - nT; T]$.

No of Nodes Out ($T$): The sum of nodes that left the bounds of $A$ during at interval $[T - nT; T]$.
No of Nodes ($T - nT$): The sum of nodes in the bounds of $A$ at time t-nt.
No of Nodes ($T$): The sum of nodes in the bounds of $A$ at time $t$.
$\alpha$: a positive value between 0 and 1 represent in advance.

After computing the mobility of all individual nodes in the network, choose the node with minimum mobility value $M_A(t)$ while forming the node clustering, It will increase the network stability [10].

Afterwards calculating the individual node mobility, we can represent the network mobility in a constant time periods.

Average node mobility of a network can be calculated by using the equation

$$\mathrm{MoB}(t) = \frac{1}{N} \sum_{n=0}^{N-1} M_n(t) \tag{5}$$

where, $N$ is that the sum of nodes within the network.

The average mobility value is less, and then the network is measured as stable network.

In extension we can determine the average time of the network mobility during the simulation period ($T$)

$$M(N) == \frac{nr}{T} \sum \mathrm{MoB}(t) \tag{6}$$

where, $K \in \{0, t, 2t, T\}$. $T$ is the time of simulation.

## *4.3 Average Distance Between the Neighbor Nodes*

The Time of Arrival (TOA) and Time Difference of Arrival (TDoA) are the two ranging methods used for calculating distance between two mobile nodes as well as position of mobile nodes.

### 4.3.1 The Time of Arrival (TOA)

The Time of Arrival (TOA) is a common ranging techniques used in the Global Positioning system (GPS). This method is based on three factors such as (a) the accurate time that a message was sent from the source node. (b) The accurate time the message arrives at a reference node. (c) The Speed at which the message travels. If these factors are known the distance between two nodes are computed by the equation.

$$D = S * (T_{\text{received}} - T_{\text{sent}}) \tag{7}$$

where $S$ is the speed of the message using this distance the set of possible locations of the source node can be identified.

### 4.3.2 Time Differences of Arrival (TDoA)

The Time Differences of Arrival is another popular ranging techniques and it uses a dynamic approach. This method does not require the time that the message was sent from the source node. This method requires only the time message was received and the speed that the message travels. Once the message is received at two neighbor hub nodes, the difference in arrival time can be used to calculate the difference in distance between the source node and the two neighbor hub nodes.

This difference can be calculated using the equation

$$\Delta D = S * (\Delta T) \tag{8}$$

where $S$ is the speed of message and $\Delta T$ is the difference in arrival times at each reference node.

## *4.4 Transmission Delay*

The transmissions delay can be computed based on the bit rate ($R$) and the length of the data packets ($L$)

$$\text{TD} = L/R \tag{9}$$

where TD is the Transmission Delay, $L$ is The Length of the data packets, $R$ is the bit rate.

Assuming that the data packets are delivered as first-come-first-serve basis as a general method used in packet switching networks; our packets could be delivered particularly after all the data packets that have reached before it has been delivered. Represent the length of the data packets by $L$ bits, and signify the transmissions rate of the connection channel from router $A$ to router $B$ by R bits/s. In case for a 100 Mbps Ethernet connection link, the rate is $R = 100$ Mbps; for a 200 Mbps Ethernet connection channel, the rate is $R = 200$ Mbps. The delay in transmission (DT) is $L/R$. This is the total time required to transmit entire data packets bits into the connection link. The delays in transmission (DT) are measured generally in the form of microseconds to milliseconds.

# 5 Proposed Algorithm

## 5.1 Node Quality Based Cluster Head Selection Algorithm

Let $M(X) = \{X_1, X_2, X_3, \ldots, X_n\}$ be the set of Nodes in the Mobile Networks and $M(V) = \{V_1, V_2, \ldots, V_k\}$ be the set of Possible Cluster Head nodes.

**Inputs**: Set of mobile nodes $G(V, E)$, distances, and neighborhood are inputs

**Output**: Set of cluster heads are output

1. Define number of clusters $k$ in a network
2. Initialize Cluster Head (CH) by first shuffling the set of mobile nodes and then randomly Selecting $K$ mobile nodes for the Cluster Head (CH) without replacement.
3. Processing the data until there is no change to the Cluster Head (CH)
    3.1. Compute the sum of the distance between mobile nodes and all Cluster Head (CH)
    3.2. Assign each mobile nodes to the nearest Cluster Head (CH)
    3.3. Compute the Cluster Head (CH) for the cluster by taking the average of the all mobile nodes that belong to each cluster.
4. Compute the energy consumption using

$$D(V_i) = \sum\nolimits_{j=1}^{n} \mathrm{dis}(v_i\ v_j)$$

5. Calculate the distance between each mobile node and cluster Head.
6. Assign the data points to the cluster center whose distance from the cluster center is minimum of all cluster centers.
7. Compute the Mobility of a Node by using the formula

$$\mathrm{MoB}(t) = \frac{1}{N}\sum_{n=0}^{N-1} MN(t)$$

8. Compute the transmission delay between the cluster head node and the intermediate node using the formula TD = $L/R$
9. Compute the residual battery energy and Compute the combined weight $W(V_i)$

$$W(V_i) = w_1.A(V_i) + w_2.E(V_i) + w_3.M(V_i) + w_4.\mathrm{TD}(V_i)$$

The combined weight $W(V_i)$ is called as Sustainable Cell rate (SC)

10. Sort $W(V_i)$ in the increment form.
11. If $W(V_i)$ is not null
12. Start $V_i \leftarrow W(V_i)$

**Table 1** Node priority model

| Node type ($v_i$) | Set name | Condition |
|---|---|---|
| 1 | St($v_i$): Strong node | Deg($v_i$) >= 3 |
| 2 | Wk($v_i$): Weak node | Deg($v_i$) >= 2 |
| 3 | Br($v_i$): Border node | Deg($v_i$) >= 1 |
| 4 | Is($v_i$): Isolated node | Deg($v_i$) >= 0 |

13. CH ← $V_i$
14. Transmit the data packets
15. End

## 5.2 *Algorithm for Removing Malicious Nodes from the MANET*

**CH**: Cluster Head **MN**: Malicious Nodes

1. Begin
2. Each CH monitor the nodes behavior and prepare the MN list
3. Transmit this list to all nodes in the network
4. After getting the malicious list information, all nodes in the network find the NodeID of the malicious nodes in their table.
5. Each node deletes all the entries related to these NodeID from the corresponding tables.
6. End

(a) Node Priority Model:

Set Node Priority table displays that strong nodes priority > weak nodes priority > border nodes priority > isolated nodes priority.

(b) *Node Quality Measurement:* Based on the type of node and its degree of connectivity the node quality can be measured.

$$\begin{aligned}\text{Node type}(v_i) = \{1, \text{Deg}(v_i) &>= 3\\ 2, \text{Deg}(v_i) &>= 2\\ 3, \text{Deg}(v_i) &>= 1\\ 4, \text{Deg}(v_i) &>= 0\}\end{aligned}$$

The Quality of the node $v_i$ is calculated directly with this equation

$$\text{Nq}(v_i) = \text{node type}(v_i) * \text{Deg}(v_i) \tag{11}$$

**Table 2** Simulation parameters

| Parameter | Value |
|---|---|
| Simulator | GloMoSim v2.03 |
| Antenna model | Omni antenna |
| Application type | Constant bit rate (CBR) |
| Data rate | 5 packet/s |
| MAC layer | IEEE 802.11, CSMA |
| Mobility model | Random way point |
| Mobility speed | 10–50 m/s |
| Number of mobile nodes | 50, 100, 150, 200, 250, 300 nodes |
| Packet size | 512 bytes |
| Pause time | 30 s |
| Propagation model | Two ray ground |
| Routing protocols | AODV, DSR, and WRP |
| Simulation area | 200 * 200 |
| Simulation time | 1000 s |
| Transmission range | 250 m |

## 6 Simulation Environment

In this part we assess the performance of our algorithm by simulations by using various simulation parameters; we evaluate our proposed algorithm with existing Cluster Head selection algorithm such as Cluster Head selection Fuzzy Logic (CHFL), Linked Clustering Algorithm (LCA) and Adaptive Clustering Algorithm. Following table gives simulation parameters (Table 2).

*Performance metrics:* There is several performance measurements used for evaluating the performance of a network. The factors are Packet transmission rate, transmission delay, energy consumptions of the network and residual energy.

## 7 Simulation Results and Analysis

**Throughput**

**Energy consumption**

**Residual Energy**

**Transmission Delay**

**Malicious Nodes Detection Ratio**

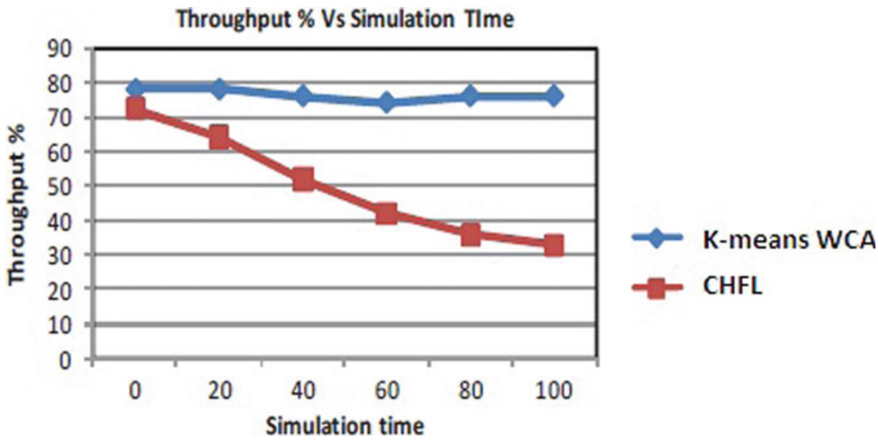



**Fig. 3** Throughput versus simulation time

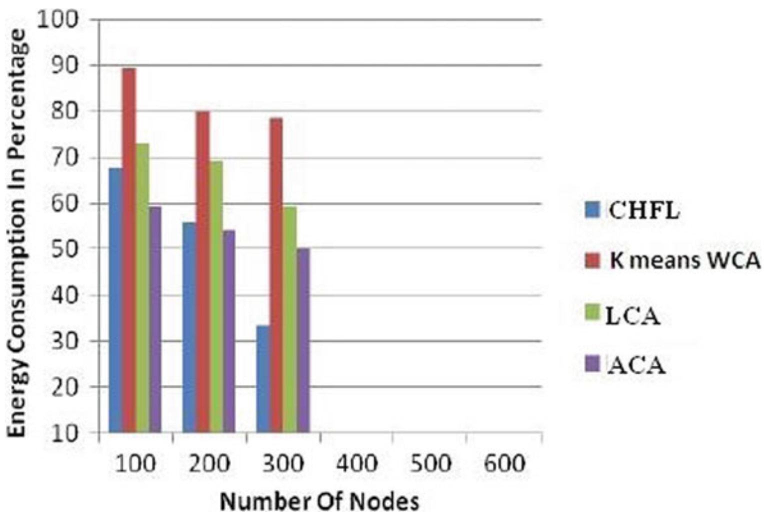



**Fig. 4** Energy consumption

The proposed system performance is analyzed with quality measurement factors of wireless sensor networks such as throughput, energy consumption, Residual energy level, packet transmission time and malicious node detection efficiency.

In Fig. 3 shows the results of proposed system throughput. It is also identify that the value of throughput and node quantity is inversely proportional with one other as a result of the number of mobile nodes increased formerly the number of data packets is added in the network which initiate network traffic to decreasing network throughput in the both cases. The proposed K means WCA algorithm is compared

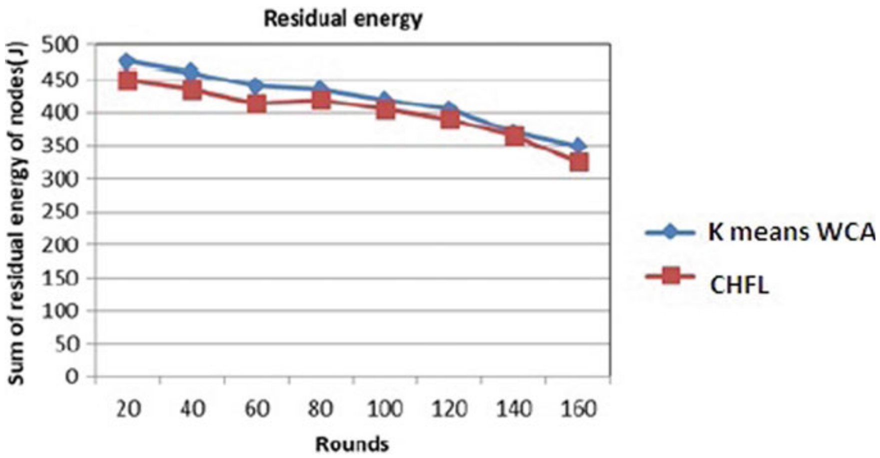



**Fig. 5** Total residual energy

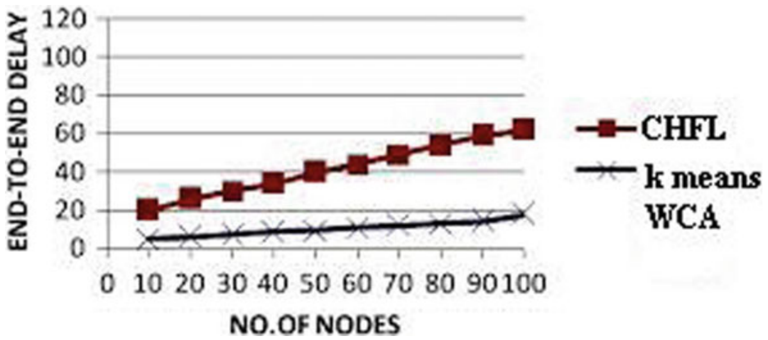



**Fig. 6** Transmission delay analysis

with Cluster Head selection using Fuzzy Logic, the CHFL gives throughput as 75% and proposed algorithm gives 80% with the simulation running time is 100 s.

In Fig. 4 gives the results of energy consumption analysis with the existing clustering algorithms, the graph shows clearly that proposed algorithm utilize 88% of energy when 100 nodes are used, 80% of power consumption when the network is formed by using 200 nodes and 78% of power is consumed when network is formed with 300 nodes. The power consumption is considerably less when compared with existing algorithms.

Figure 5 shows that the remaining battery power of nodes at different periods of time in fixed nodes speed of 20 m/s. The proposed algorithm is having high residual energy when compared with existing algorithm. The proposed method confirms to be more energy aware system.

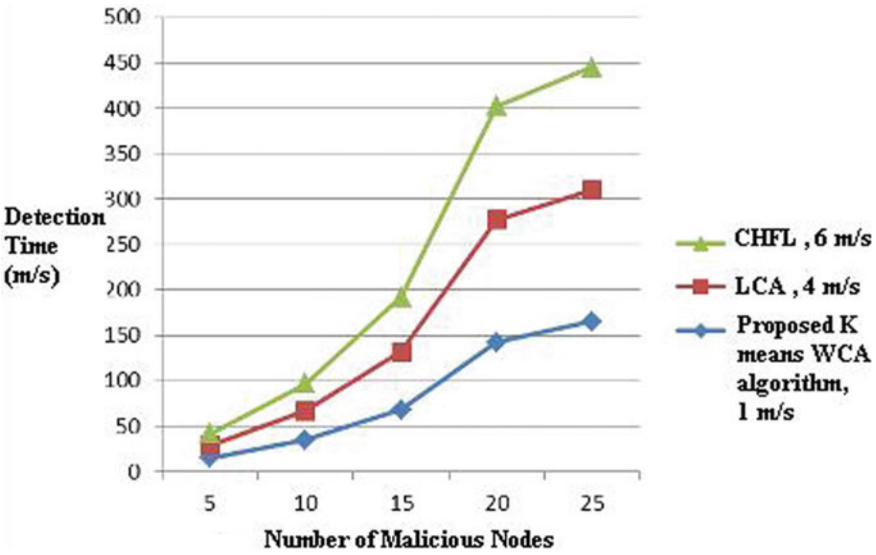



**Fig. 7** Analysis of malicious node detection efficiency

Figure 6 shows the results of transmission delay of proposed algorithm with CHFL algorithm. The proposed algorithm having maximum 20% delay when the network uses 100 nodes. A secure route without suspicious nodes for data transmissions reduces the latency.

Figure 7 shows that the comparison of malicious node detection efficiency in a various tine interval with existing algorithms such as Cluster Head Fuzzy logic (CHFL), Linked Clustering Algorithm (LCA) and proposed K means WCA algorithm. The CHFL takes average detection time as 6 m/s, LCA having average detection time as 4 m/s and proposed algorithm takes 1 m/s as average malicious nodes detection time.

## 8 Conclusions and Future Works

In this paper, we have proposed a unique optimized K-means algorithm established with Weighted Clustering Algorithm (WCA).This algorithm compensates and decreases energy consumption by electing the efficient Cluster Head (CH) node by considering the essential mobile node clustering factors such as Node mobility, Distance between CH and base stations, Energy consumption of a node, Residual energy and link quality. All possible set of CH nodes in a wireless network must compute the values of all these factors individually. The Cluster Head (CH) nodes are ranked based on these values. A node obtain high priority value will be elected as a Cluster Head (CH) node. After a first node finishes its role, second node will be preferred from these lists so the time complexity will be reduced. Simulation a result shows that the proposed algorithm provides the optimum solution in node clustering and Cluster Head (CH) selection challenges. The simulation can be performed with other node clustering algorithms such as CHFL, LCA and ACA. The proposed algorithm effectively increases the network lifetime by selecting a efficient Cluster Head (CH) node and also increases the network throughput. In future work we propose algorithm enhancements for selecting an optimum CH node to detect and prevent new type of attacks and increasing additional selection criteria of CH node for improving the Quality of Service (QoS) of a network, perform simulation with different scenarios.

## References

1. Muruganandam J, Renjit A, Sendhil Kumar R (2019) A survey: comparative study of security methods and trust manage solutions in MANET. In: Fifth international conference on science technology engineering and mathematics. https://doi.org/10.1109/ICONSTEM.2019.8918697
2. Kumari P, Aggarwal G, Singh S (2018) Clustering in mobile Adhoc network: WCA algorithm Springer Nature Switzerland AG2019. Online ISBN: 978-3-030-03146-6. https://doi.org/10.1007/978-3-030-03146-6_58
3. Abbasi AA, Younis M (2007) A survey on clustering algorithms for wireless sensor networks. Elsevier Comput Commun 30(14–15):2826–2841. https://doi.org/10.1016/j.comcom.2007.05.024
4. Thomas L, Annappa B (2011) Application of parallel K-means clustering algorithm for prediction of optimal path in self aware mobile ad-hoc networks with link stability. In: Abraham A, Mauri JL, Buford JF, Suzuki J, Thampi SM (eds) Advances in computing and communications. ACC 2011. Communications in computer and information science, vol 193. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-22726-4_42
5. Du J, You Q, Zhang Q et al (2017) A weighted clustering algorithm based on node stability for ad hoc networks. In: 16th international conference on optical communications and networks (ICOCN). Wuzhen, China. IEEE. https://doi.org/10.1109/ICOCN.2017.8121247
6. Tay M, Senturk A (2019) Energy-aware clustering algorithms used in wireless sensor. In: Networks 2019 scientific meeting on electrical-electronics and biomedical engineering and computer science (EBBT) Istanbul, Turkey, Turkey, IEEE Xplore. https://doi.org/10.1109/EBBT.2019.8742068
7. Subathra P, Radha Krishnan P, Sivagurunathan S (2012) A cluster-based reliable token circulation scheme for group communication in MANET. Arab J Sci Eng 37:647–664. https://doi.org/10.1007/s13369-012-0195-4
8. Paliwal P, Kumar A (2015) Cluster based probabilistic approach to improve MANET performance In: 2015 international conference on soft computing techniques and implementations (ICSCTI), Faridabad, India. IEEE. https://doi.org/10.1109/ICSCTI.2015.7489604
9. Jyothi R, Cholli NG (2019) New approach to secure cluster heads in wireless sensor networks. In: 2019 5th international conference on advanced computing &communication systems (ICACCS), Coimbatore, India. IEEE. https://doi.org/10.1109/ICACCS.2019.8728380
10. Renjit JA, Shunmuganathan KL (2010) Implementing advanced intrusion detection system by monitoring network anomalies and using encrypted access of data. Natl J Syst Inform Technol 3(1):32–41. ISSN 0974-3308. JUNE 2010©SRIMCA, New Delhi
11. Kukreja D, Dhurandher SK, Reddy BVR (2018) Power aware malicious nodes detection for securing MANETs against packet forwarding misbehavior attack. J Ambient Intell Human Comput 9:941–956. https://doi.org/10.1007/s12652-017-0496-2

12. Aranganathan A, Suriyakala CD (2010) An efficient secure detection and prevention of malevolent nodes with lightweight surprise check scheme using trusted mobile agents in mobile ad-hoc networks. J Ambient Intell Human Comput 10:3493–3503. https://doi.org/10.1007/s12652-018-1069-8
13. Singh K, Verma AK (2020) TBCS: a trust based clustering scheme for secure communication in flying ad-hoc networks. Wirel Pers Commun. https://doi.org/10.1007/s11277-020-07523-8
14. Sharma DK, Dhurandher SK, Agarwal D et al (2019) kROp: k-means clustering based routing protocol for opportunistic networks. J Ambient Intell Human Comput 10:1289–1306. https://doi.org/10.1007/s12652-018-0697-3
15. Chatterjee M, Das SK, Turgut D (2020) WCA: a weighted clustering algorithm for mobile ad hoc networks. Cluster Comput 5:193–204. https://doi.org/10.1023/A:1013941929408
16. Vijayan R, Jeyanthi N (2019) Trust management approaches in mobile adhoc networks. In: Jeyanthi N, Abraham A, Mcheick H (eds) Ubiquitous computing and computing security of IoT. Studies in Big Data, vol 47. Springer, Cham. https://doi.org/10.1007/978-3-030-01566-4_4
17. Vatambeti R (2020) A novel wolf based trust accumulation approach for preventing the malicious activities in mobile ad hoc network. Wirel Pers Commun 113:2141–2166. https://doi.org/10.1007/s11277-020-07316-z
18. Josephine A, Senthilkumar J, Tanimoto S (2020) Support vector regressive linear program boost based node trust evaluation for secure communication in MANET. Wireless Pers Commun. https://doi.org/10.1007/s11277-020-07209-1
19. Guaya-Delgado L, Pallarès-Segarra E, Mezher AM et al (2019) A novel dynamic reputation-based source routing protocol for mobile ad hoc networks. J Wireless Com Network 2019:77. https://doi.org/10.1186/s13638-019-1375-7
20. Almazyad AS (2018) Reputation-based mechanisms to avoid misbehaving nodes in ad hoc and wireless sensor networks. Neural Comput Appl 29:597–607. https://doi.org/10.1007/s00521-016-2555-6
21. Menaka R, Ranganathan V, Sowmya B (2017) Improving performance through reputation based routing protocol for manet. Wireless Pers Commun 94:2275–2290. https://doi.org/10.1007/s11277-016-3396-4
22. Sapna, Kulkarni B, Yuvaraju BN (2018) Trust value updation algorithm for multicast routing algorithm for cluster based MANET. In: 2017 international conference on wireless communications, signal processing and networking (WiSPNET). https://doi.org/10.1109/WiSPNET.2017.8299962

# Prediction of Liver Patients Using Machine Learning Algorithms

**Shefai Tanvir Fayaz, G. S. Tejanmayi, Yerramasetti Kanaka Ruthvi, S. Vijaya Shetty, Sharada U. Shenoy, and Guruprasad Bhat**

## 1 Introduction

Liver is a massive organ present in the human body that lies on the right side of the abdomen, under the rib cage and above the stomach. The weight of the liver is approximately three pounds and is seen to be reddish-brown color. The organ is as large as the size of a football and is very essential for digesting food and excreting the toxic substances from the body which, if not carried in a regular process, can lead to hazardous complication. There are two large sections in the liver, called the left and the right lobes. The gallbladder sits under the liver, alongside pancreas and intestines. The organs around and along with the liver absorb, digest, and process food. Liver predominantly filters the blood running in the digestive tract, just before the blood is passed to the rest of the body, this being the most imperative activity. Another substantial process is the detoxification of chemicals and metabolism of drugs. Secretion of bile is carried on by the liver that ends up in the intestines as it detoxifies. It is also responsible in making proteins significant for blood clotting and additional functions [1]. Harm of the liver is one of the deadliest illnesses on the planet. The vital carter of harming the liver is fatty liver, liver fibrosis, cirrhosis, hepatitis, and other diseases. It is remarkably hard to identify damages at the beginning of liver illness, even though liver tissue has already been harmed. To analyze

S. Tanvir Fayaz (✉) · G. S. Tejanmayi · Y. Kanaka Ruthvi · S. Vijaya Shetty · S. U. Shenoy
Department of Computer Science and Engineering, NMIT, Bengaluru 560064, India

S. Vijaya Shetty
e-mail: vijayashetty.s@nmit.ac.in

S. U. Shenoy
e-mail: sharadauday@nitte.edu.in

G. Bhat
CISCO Systems Pvt. Ltd, Bengaluru 560103, India
e-mail: gurbhat@cisco.com



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_12

the damage, a good number of specialists need to offer their time. Early finding is essential to spare the patient with all the pharmaceutical medications that are given in the absence of recognition of the cause of the problems. Commonly observed liver diseases are as follows:

- Fatty liver is a condition which is revocable and where vacuoles of triglyceride fat in large size acquire in liver cells via the limit process. The occurrence of fatty liver is found in people who consume high levels of alcohol, yet in people who never had alcohol.
- Cirrhosis is another important type of liver damage. It is usually the result of long-term damage of liver. When liver is damaged for a long time and starts to malfunction, this particular type of liver damage occurs.
- Hepatitis is generally caused by an infection that extents by excess contamination or direct contact with tainted body.
- Liver cancer risk is greater on those who have cirrhosis or people with viral hepatitis. Often it spreads from other organs to liver.

## 2 Literature Survey

Previous researches for detection of patients of liver diseases have given good accuracies above 90%, be it using machine learning or data mining techniques. In [2], a three-phase analysis is employing normalization, PSO followed by application of algorithms. It showed J-48 having highest accuracy. A comparison of greedy stepwise versus PSO is also given, of which PSO serves to be the best. Papers [3] and [4] also have adopted machine learning approach, testing different algorithms, and their accuracies. This has been performed on different datasets [4]. Showed an accuracy of 92.8% using artificial neural network with 10 inputs, 2 layers, and hundreds of neurons in hidden layers. A GUI has also been created.

In [5], data mining techniques involving extract, transform, load, and data mining algorithms have been incorporated to find the class of patients, i.e., liver patients or not. A decision tree based on labels shows the same. A highest accuracy of around 67% was achieved. In [6, 7] and [8], various data mining algorithms have been used, among which genetic programming, PSO K-Star, and NB tree showed the highest accuracies, respectively. Unlike the other two papers, in [6], concentration has just been laid on Bayesian classification performed using Weka tool. This yielded an accuracy of 72.6%. In [7], fuzzy rules have been generated to get to an accuracy of 94%.

Paper [9] gives the importance of feature selection in areas as such for enhanced performance. Selecting apt features that enhance the performance of algorithms plays a crucial role so does elimination of unimportant one's that pull back the accuracy. In [10], methods have been defined using which imbalanced datasets can be handled. This includes sampling techniques based on size of dataset which can be employed keeping into consideration the pro's and con's which come with it. Oversampling is done to increase the size of data and undersampling to minimize it. Oversampling can be used for small amount of data, though this increases the processing time, it enables better learning for algorithms, whereas undersampling is better for large datasets but is not advised to be used. As it compromises the quality of data and data gets lost as well.

In [11], specific use of Bayes network worked in WEKA tool yielded an accuracy of about 73% [12]. Shows the use of C4.5 on dataset of student academic performance using confusion matrix which throws light on effectiveness of this algorithm in machine learning. Paper [13] gives liver disease classification of cirrhosis, bile duct, chronic hepatitis, liver cancer, and acute hepatitis on a generic view using Naive Bayes and SVM.

Some of the previous works have shown efficiency above 90%. Our aim is not only to achieve a better efficiency but also employ extensive preprocessing so that the obtained result is as robust as possible due to the application of over sampling. Also, we aim at providing etiological survey results which makes this project unique.

## 3 System Architecture

When machine learning is applied for any analysis, it is important to develop a solid architecture for it. System architecture refers to the underlying conceptual model of the system. A good architecture covers all crucial concerns like system views, the structure of the system, as well as how the system behaves under certain conditions. In other words, a strong grasp of the system architecture can help the user to understand the scope as well as the limitations of the said system. In machine learning, system architecture is defined as the subject that helps in the evolution of any idea from the concept of fantasy to the proof of reality. The system architecture, like that of the most large-scale structures, comprises of smaller components that will be attempted to individually explain. Some of the various parts of the system are:

- **Dataset**: One of the key principles of machine learning is the ability to infer the core trends that accede to in the data, by oversampling a comparatively small size of the data.
- **Training dataset**: The sample data that is taken is known as the training data. The training data is randomly taken from the complete dataset.
- **Testing dataset**: Testing data is a larger sample of data that is also taken from the complete dataset, and as previously explained, the accuracy is measured. Models require the use of similar data for the training and testing data. That helps in the reduction of data discrepancies, as well as better understanding the primary model of the system.
- **Normalization**: It is a common method applied as part of preprocessing in machine learning to get robust results. The objective of this is to alter the values of columns that are numeric in the dataset to a scale which is common, without misrepresenting and losing the meaning in the differences of the range in the

values. It is required when features have dissimilar ranges or there are any null values, in that case average column value replaces the null value.
- **Sampling**: Sampling benefits a lot when there is less or more of data. Sampling is a significant factor that defines the precision of any research result. The way the data is sampled will openly reflect in the final outcome. There are a lot of methods in sampling that can be used depending upon the need and situation. In our case, we have used oversampling to increase the number of records in the dataset.
- **Feature extraction**: In feature extraction, creation of combinations of variables is done, in order to get the best accuracy. Properly optimized feature extraction is the crucial factor for effective model construction. Particle swarm optimization (PSO) is the method used here. It treats every particle or a point in a swarm of particles or a dataset, as a candidate solution. That is how the computation is carried by the PSO. PSO is used in many medical applications such as artificial neural network training, classification of data, and network optimization, for its ability to improve the probability of accurate performance.
- **Classification models**: In order to draw some deduction from observed values, a classification model is used. It will try to predict the value of the outcomes with the inputs given.

  – *Decision trees:* Decision trees are a class of supervised machine learning where the data is continuously divided according to a definite parameter.
  – *Random forest:* Random forest, like its name infers, contains a large number of single decision trees that operate as an ensemble. Individual trees in random forest give a class prediction and the class with the most votes becomes the model's prediction.
  – *J48:* A decision tree is generated using the J48 algorithm. These decision trees that are generated by J48 are used for classification.
  – *SVM:* Support vector machine is a discriminative classifier which results in classification by defining a separating hyperplane. Meaning, with the given labeled training data, the algorithm produces an optimal hyperplane resulting in categorization of new examples.

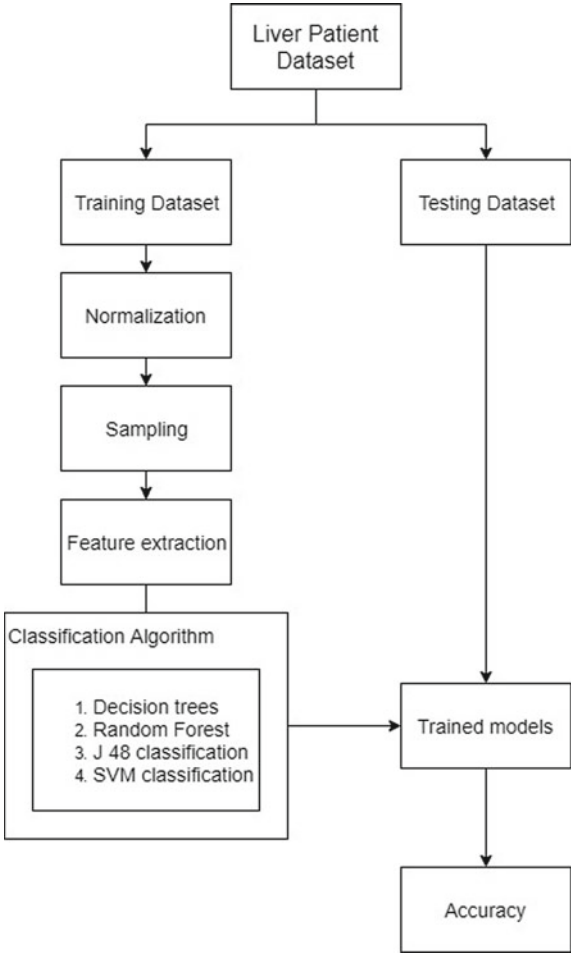
- **Accuracy**: Based upon the various algorithms, in the end, a final output of the accuracy is derived that has been predicted upon the given data.

The flow diagram in Fig. 1 depicts how the data is dealt with through the various stages of preprocessing after being split into training and testing data. Finally, machine learning algorithms are applied onto data models to get the accuracies of each.

## 4 Methodology

For the implementation, various methods, for the analysis of the data, have been applied to get better and more precise results. As mentioned previously, extensive

**Fig. 1** Architectural design




preprocessing is done. By extensive preprocessing, we mean good amount of preprocessing, involving normalization, oversampling, and PSO. This gives us a better hand at accuracies and etiological survey seen in the results chapter.

During the whole course of implementation, the approach has been pragmatic rather than justifying results solely based off assumptions. The usefulness of each algorithm will be proved by their accuracy that will be specified in further sections. On a generic view, the project can be viewed as:

- **Data**: For the collection of data, we have used the standard dataset that was generated by the UCI machine learning repository. For liver patient prediction, a dataset with properties like albumin, bilirubin levels, and some others were required. The dataset obtained fulfills all the qualities needed for the project. However, some minor adjustments had to be set regardless such as cleaning these from typing errors and other small mistakes that would create uncertainty in the final result.

- **Language and environment**: The project has been carried out in Python language, and this includes actual algorithms. Python is a dynamic language that was built relatively recently compared to other programming languages such as C and Java. Hence, Python tends to fix a lot of problems that were encountered by the users of said languages. Python has an in-built garbage collector as well as support for Unicode. Python emphasized white space, enhancing code readability as well as an increased likelihood of large-scale projects to be built due to the same. Python also contains various libraries that are specifically designed to help in the generation of machine learning projects such as sci-kit learn, which is used in this project. In addition to this, there are Python interpreters such as Anaconda, which supports other interpreters such as Spyder and the Jupyter notebook. It is due to all of these reasons, from the general Python language, to the various libraries, along with the various interpreters, that caused the choosing of this language over several others.
- **Algorithms**: The algorithms used were standard machine learning algorithms whose use and general execution have been explained in depth.

## 5 Results

Chosen algorithms have been applied onto the dataset with various kinds of preprocessing which have later been combined as per their compatibility. This shows how the accuracy is either enhanced or reduced. Below are two tables depicting the same without and with application of PSO, respectively.

As seen in Table 1, we have the dataset that has been preprocessed with normalization, oversampling, and a combination of oversampling and normalization. When the dataset is introduced to the SVM classifier, random forest and J48 classifier without any preprocessing, J48 yields a decent result of 87.5%, which is the highest so far. SVM gives us its best result of 78% accuracy with oversampled and normalized data, suggesting that the disadvantage being lack of data, although 78% is notably lesser against 87.5% of J48. Likewise, 85.5% is the best accuracy given by random forest with normalized and oversampled data. Random forest being closer to J48 but falling behind by 2% is pointedly far from SVM.

**Table 1** Accuracy percentages of algorithms without feature selection

| Algorithms\preprocessing | None | Normalization | Oversampling | Oversampled and normalized | Normalized and oversampled |
|---|---|---|---|---|---|
| SVM | 71 | 68 | 69 | 78 | 71 |
| Random forest | 71.8 | 65.8 | 82.5 | 75.3 | 85.5 |
| J48 | 87.5 | 87.5 | 80.1 | 74.7 | 84.9 |

**Table 2** Accuracy percentages of algorithms with feature selection

| Algorithms\preprocessing | PSO | PSO and normalization | PSO and oversampling | PSO and oversampled and normalized | PSO and normalized and oversampled |
|---|---|---|---|---|---|
| SVM | 68 | 80 | 68 | 71 | 72 |
| Random forest | 69.2 | 70.1 | 82.5 | 75.3 | 85.5 |
| J48 | 84.7 | 81.8 | 95.6 | 95.6 | 94.7 |

Since the highest recorded accuracy was 87.5%, obtained from J48 with normalization and dataset with no preprocessing, we introduce the dataset to PSO. PSO is an optimization technique used in preprocessing of data which leads to promising results. Table 2 illustrates the various accuracies obtained from introducing PSO to dataset after applying the previous preprocessing techniques as shown in Table 1. Here, SVM classifier yields its good result of 80% accuracy with normalization as opposed to 78% from Table 1. For random forest, 85.5% accuracy remains the same compared to Table 1. It can be said on a safer note that random forest was not affected with the introduction of PSO. J48 tops the accuracy list with 95.6% for both oversampled and, oversampled and normalized data after the application of PSO to dataset. Hence, PSO did show its capabilities of yielding good results with SVM and J48.

Figure 2 shows the various attributes present in the dataset of Indian Liver Patients namely age, gender, tot proteins (total proteins), albumin (albumin), tot bilirubin (total bilirubin), direct bilirubin (direct bilirubin), ag ratio (albumin and globulin ratio), sgpt (alamine aminotransferase), sgot (aspartate aminotransferase), alkphos (alkaline phosphotase) and is patient (liver patient or not) represented in the subimages (a), (b), (c), respectively. Blue areas in the image represent the patients suffering from liver diseases, and red area represents the non-liver patients. There are 416 patients suffering from liver diseases and 167 of non-liver patients.

Age ranges from 4 to 90 as any patient aged above 89 is recorded as 90 in the dataset. As we can see, the blue peak near the range of age 47 is highest, suggesting that the age around 47 is vulnerable for liver diseases. Gender, we have 441 male and 142 female patient records. A lot of blue area in the male category suggests that the male are more susceptible for liver diseases. Bilirubin is a yellowish material which is formed by the breakage of red blood cells. The normal total bilirubin levels are between 0.3 and 1.2 mg per deciliter (mg/dL), whereas direct bilirubin levels are generally between 0 to 0.4 mg/dL. Liver is responsible to generate most of our proteins in the body. Two of the important proteins are albumin and globulin. The total protein value is the total of both of these proteins in our bodies and the normal range of this is between 6 and 8.3 g per decilitre (g/dL). Albumin normal range alone is 3.4–5.4 g/dL. Aminotransferases are the sensitive and largely used enzymes in the body. In case of any damage to the liver, the levels of these enzymes increase. The normal range for the SGPT (alamine aminotransferase) and SGOT (aspartate aminotransferase) is 7–56 units per liter of serum and 5–40 units per liter of serum,

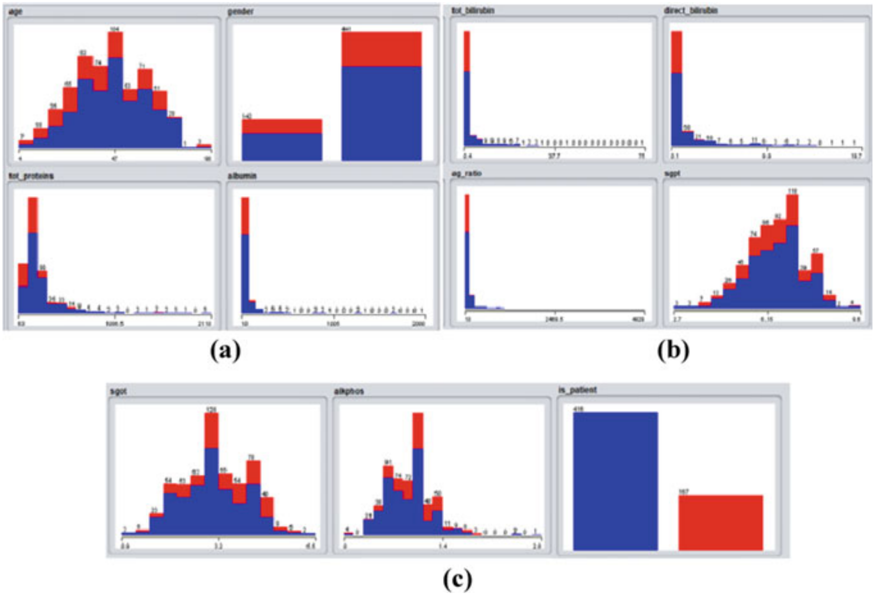



**Fig. 2** Graphical representation of the attributes present in the Indian Liver Patients Dataset obtained from UCI machine learning repository. The attributes are namely age, gender, total proteins, and albumin as shown in **a**, total bilirubin, direct bilirubin, ag ratio, and SGPT as shown in **b**, and SGOT, alkaline phosphate, and is_patient in **c**

respectively. Alkaline phosphotase (alkphos) is other type of enzyme whose normal range is 20–140 units per liter of serum. These values vary from person to person depending on their age, gender, etc.

Figure 3 gives comparative study of algorithm accuracies in graphical form. Accuracies of algorithms without any preprocessing versus that of various preprocessing are depicted. It shows how extensive preprocessing done has improved algorithm accuracies except for that of SVM, which shows different behavior.

The values shown in Table 3 were given to the J48 model with PSO applied on oversampled and normalized data which has the best accuracy. The model predicted it correctly, as a "not patient" and a "patient," respectively.

A comparison of actual values and the values predicted by J48 has been given in Table 4, of which eight records have been classified correctly out of nine.

## 6 Conclusion

In this paper, a complete discussion of the ML techniques and its applications like disease prediction or prognosis is presented. The study proposed focuses on the development of predictive models using supervised machine learning methods and

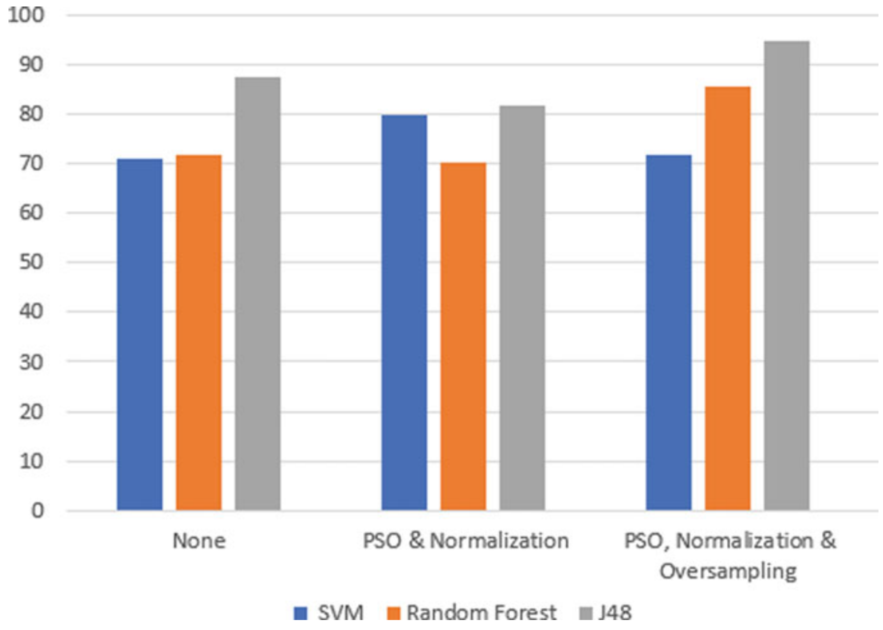



**Fig. 3** Graphical representation of algorithm accuracies

**Table 3** Test records of a patient suffering from liver disease and a non-patient

| Age | Gender | Tot_bilirubin | Direct_bilirubin | Tot_proteins |
|---|---|---|---|---|
| 17 | Male | 0.9 | 0.3 | 202 |
| Albumin | ag_ratio | sgpt | sgot | alkphos |
| 22 | 19 | 7.4 | 4.1 | 1.2 |

**Table 4** Few sets of actual and predicted values of patients

| Actual values | Predicted values |
|---|---|
| Not_patient | Not_patient |
| Patient | Patient |
| Patient | Patient |
| Patient | Patient |
| Not_patient | Patient |
| Patient | Patient |
| Patient | Patient |
| Not_patient | Not_patient |
| Patient | Patient |

classification algorithms aiming to predict valid disease outcomes. The amalgamation of heterogeneous data with the different techniques of feature extraction and classification can prove to be efficient tools in the medical domain. This dataset is analyzed to predict the occurrence of liver diseases using ML algorithms such as J48, decision trees, SVM, and random forest.

These algorithms give distinguishable results based on particle swarm optimization (PSO) feature selection model. It is important that the data is examined continuously and categorically. PSO solves a variety of difficult optimization problems, and it has faster convergence rate than other algorithms of feature selection. And, PSO is an outstanding robust evolutionary strategy. Oversampling is capable of improving resolution and signal-to-noise ratio and normalization helps to achieve a more linear and robust relationship. When PSO is combined with these two preprocessing steps, namely oversampling and normalization, significant results are observed from the algorithms. On the flip side, SVM gave an accuracy of 80% with PSO and normalization, being it is highest. J48 and random forest classifications with particle swarm optimization (PSO), oversampling, and normalization gave accuracies of 95.6% and 85.5%, respectively. Thus, with extensive preprocessing, highest accuracy of 95.6% with J48 is achieved.

## References


1. Johns Hopkins Medicines, 5 Reasons You May Be at Risk for Liver Disease. https://www.hopkinsmedicine.org/health/wellness-andprevention/5-reasons-you-may-be-at-risk-for-liverdisease
2. Banu Priya M, Laura Juliet P, Tamilselvi PR (2018) Performance analysis of liver disease prediction using machine learning algorithms. Int Res J Eng Technol (IRJET) 5(1)
3. Jacob J, Mathew JC, Mathew J, Issac E (2018). Diagnosis of liver disease using machine learning techniques. Int Res J Eng Technol (IRJET) 5(4)
4. Sontakke S, Lohokare J, Dani R (2017). Diagnosis of liver diseases using machine learning. In: International conference on emerging trends and innovation in ICT (ICEI)
5. Rajeswari P, Sophia Reena G (2010) Analysis of liver disorder using data mining algorithm. Glob J Comput Sci Technol 10(14) (Ver. 1.0)
6. Baitharua TR, Panib SK (2016) Analysis of data mining techniques for healthcare decision support system using liver disorder dataset. In: International conference on computational modeling and security
7. Aneeshkumar AS, Venkateswaran CJ (2015) A novel approach for liver disorder classification using data mining techniques. Eng Sci Int J (ESIJ) 2(1)
8. Alfisahrin SNN, Mantoro T Data mining techniques for optimatization of liver disease classification
9. Dash M, Liu H (1997) Feature selection for classification. Intell Data Anal
10. Sonak A, Patankar RA (2015) A survey on methods to handle imbalance dataset. Int J Comput Sci Mob Comput 4(11)
11. Dhamodharan S (2014) Liver disease prediction using bayesian classification. In: Special issue, 4th national conference on advanced computing, applications and technologies
12. Budiman E, Haviluddin, Dengan N, Kridalaksana AH, Wati M, Purnawansyah Performance of decision tree C4.5 algorithmin student academic evaluation

13. Vijayarani S, Dhayanand S (2015) Liver disease prediction using SVM and Näıve Bayes algorithms. Int J Sci Eng Technol Res (IJSETR) 4(4)
14. Davis CP, Shiel Jr WC (2018) Liver blood tests abnormal values. MedicineNet

# Development of Security Performance and Comparative Analyses Process for Big Data in Cloud

**M. R. Shrihari, T. N. Manjunath, R. A. Archana, and Ravindra S. Hegadi**

## 1 Introduction

Big data is a name related among the dataset where volume is huge away from the storage and investigates competence of conventional database organization resolution. A big data is a grouping of 3 Vs. Volume is a term associated to massive size of data. Through the entrance of big data, extraordinary modify in the configure of information has to be experimental kind to rise variety of big data. Data imminent as of dissimilar organization is liable to appearance variety of information. Consequently, currently with structured data, semi-structured and unstructured data are also experiential. Structured data is well for tabular configuration, in row and column format. Therefore, conventional organization method can be without difficulty development of these categories of data but semi-structured and unstructured data are real confront for conventional database administrative tools. Community media texts are the instance of semi-structured data while audio, video, and an image approach beneath the kind of unstructured data. Current advance in big data cloud security contains lots of opportunity for researchers on cryptography to progress the security by means of authentication and authorization for right to use stored information in a competent way. Cloud computing prevents the constraint to preserve computing hardware, storage, and software. The planned scheme is used to make safe the data in big data cloud and conquer a variety of challenges related to security. A lot of security services such as authentication, privacy, and data veracity make

M. R. Shrihari (✉) · T. N. Manjunath
Department of ISE, BMS Institute of Technology and Management, VTU, Bengaluru, Karnataka, India

R. A. Archana
Department of CSE SVIT, VTU, Bengaluru, Karnataka, India

R. S. Hegadi
Department of Computer Science, Central University of Karnataka, Kalaburagi, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_13

sure by the relevant planned method. Only the authorised user is permitted to access cloud data using this technology [1].

## 2 Machine Learning Impact on Big Data

The name big data was primary invent in 1990 covenant with the learning of huge datasets. Information storage, information capturing, information study, information reservation, information revelation, and information move are a number of the confront of big data. It can perform wonder merely if the majority significant data can be extracted during it. To organize to dig expensive data starting the massive reach your peak of big data use of analytical analytics, customer performance and other big data are in development. Big data can avoid disease, identify fault, and assist in industry, financial services, etc. Machine learning is used to discover the hidden prototype from huge and composite data. Machine learning is a system during which form is educated to discover since statistics and consequently it is extensively used in approximately each field in discovery a precious prototype from big data. This method does not involve individual disturbance for generate consequence. Big data is important only if valuable data is composed since it assists suitable machine learning algorithm [2].

## 3 Security Challenges in Big Data Cloud Computing

Security in the cloud is talented, to an inadequate amount, during stranger organize and assertion greatly like in traditional outsourcing courses of action. In any case, since there is no common big data cloud computing safety standard, there are additional complexity associated to this. Numerous cloud vendors actualize their preventive measures and security novelty and execute varying security form, which must evaluate their benefits. In a seller cloud form, it is at most recent along to receiving customer organizations to assurance that safety in the big data cloud convene their own safety policies during requirements gathering, supplier chance appraisals, due persistence, and affirmation exercises. In this way, the security challenges confronted by organizations wishing to utilize big data cloud services are not drastically the same as those reliant on their in-house overseas enterprises. The same inside and outside dangers are available and require chance mitigation or hazard acknowledgment [3].

### 3.1 There Are Several Challenges in Big Data

i. The first challenge for institute is to decide and pick the significant and essential data. With such high quantity of information, it becomes essential for institute to able to detach the important data.
ii. The second challenge is that smooth now, in institute, many data points are not associated. This predicament of connectivity is a strict impediment. Big data is all concerning compilation of data as of a variety of operation position. Institute must to be able to supervise data as of transversely its venture.
iii. To influence big data, one has to occupation transversely responsibility such as information technology, engineering, and finance. Consequently the possession and procurement of this statistics have to be a supportive endeavor transversely these departments. This demonstrates to be an important executive confront.
iv. There is a security perspective associated to big data compilation. This is a most important difficulty avoid corporation since enchanting filled benefit of big data assessment. A number of concerns are to be concentrate on to confine the complete prospective of big data. Access to data is critical in companies that will increasingly need to integrate information from multiple data sources, often from third parties, and the incentives have to be in place to enable this [4].

The protection challenge in big data cloud computing atmosphere decreases over several intensity: the association intensity which incorporates commerce with association practice and association protection such as disseminated nodes, dispersed data, and transportation between the nodes; authentication intensity where the client holds encryption or decryption organism, authentication methods such as convention administrative rights, endorsement of relevance and nodes, and classification entry; the information level which is worried with data reliability and accessibility as well as data safety and data sharing. Cloud computing pursues the procedure of public possessions, wherever the confidentiality of data is extremely significant since it appears in several confront like reliability and authoritative admission. Data reliability guarantees that data is not degraded or interfere with through message. Authoritative admissions avoid data from permeation attack while backups and replicas allow access to data efficiently even in case of technological inaccuracy or adversity in a quantity of cloud location. Big data appears as a quantity of dispute as the customer can be organized into cluster datasets, dispensation, and organization argument. While industry through big quantity of data, the argument countenances are volume, variety, velocity, veracity, and verification which are also known as 5 Vs of big data. Through the factors and argument that concern the dispensation of big data in an appropriate method is the bandwidth and latency. A number of challenges are recapitulated at this time in the association between big data and cloud computing [5].

### 3.2 Data Storage

In the storage of big data, conventional storage is problematical since solid take often be unsuccessful, data safety system is not efficient, and the velocity of big data involves storage organization in order to develop quickly, which is hard to complete with conservative storage scheme [6].

### 3.3 Variety of Data

Big data logically develops, enhances, and differs, which is the consequence of the development of approximately indefinite foundation of information. The big data encompasses contrary outline and is incoherent. A client can accumulate data in structured, semi-structured, or unstructured system. A structured data system is appropriate for record scheme, whereas semi-structured data systems are solitary literally appropriate, and unstructured data is improper since it encloses a complex system that is hard to signify in rows and columns [6].

### 3.4 Data Relocate

In the data during numerous stages: data gathering, input, dispensation, and output, big data relocate is an argument, so data density method requires to be concentrated to decrease the volume, anywhere data amount is a complexity to relocate velocity [6].

### 3.5 Privacy and Data Ownership

The cloud environment is an open circumstance, and the responsibility of client to examine is incomplete. Privacy and protection are an essential argument for big data. According to International Data Corporation (IDC) estimates, by 2020, approximately 40% of comprehensive data determination can be admitted by cloud computing. Since such, there is a tough command to examine the confidentiality of in sequence and safety challenges in both cloud computing and big data. Resolve safety and secrecy challenges connected with big data and cloud computing knowledge, addressing of three problems as listed beneath, are highlighted [7].

## 4 Security Assessment of Four Levels Security Scheme

Security investigations clarify the proposed system of safe and four levels security system. The four levels of security systems namely authentication, authorization, encryption, and decryption are utilized to protect the data from unauthorized person. Further, confidentiality, integrity, efficiency, and intelligibility. Table 2 represents the dissimilar type of security examination handled by the planned work. Throughout the data admittance, registration, authentication, authorization are possible for data supplier, data user, and cloud service provider. Through the data admission, unauthenticated or unauthorized person's only inventory stage is achievable, other security segments are not potential [13]. Encryption is possible for data provider and cloud service provider (Table 1).

Decryption is potential for data user and cloud service provider. An unauthenticated or unauthorized person cannot study or download the data, so confidentiality is handling by the proposed work. Reliability is potential because data supplier can edit the file earlier than uploading it. Subsequent to upload the encrypted data, no one might edit the file. The planned work creates efficient consequence though uploading and downloading for data supplier (DS), data user (DU), and CSP. Uploading or downloading information is observable to cloud service provider which shows intelligibility. Consequently that cloud service provider might obstruct that the illegal and

**Table 1** Security services of enhanced protection scheme for big data clouds

| | Data supplier | Data user | Cloud service provider |
|---|---|---|---|
| Authentication | ✓ | ✓ | ✓ |
| Authorization | ✓ | ✓ | ✓ |
| Encryption | ✓ | × | ✓ |
| Decryption | × | ✓ | ✓ |
| OTP | × | ✓ | ✓ |

**Table 2** Role and operation function of enhanced security system for big data clouds

| Role | Operation functions |
|---|---|
| Data supplier | Encrypt data |
| | Upload data |
| Data user | Download data |
| | Decrypt data |
| Cloud service provider | Authentication—verifying username and password |
| | Authorization—verifying credentials of the user |
| | Key generation, OTP generation, PoW generation |
| | Block unauthorized user |

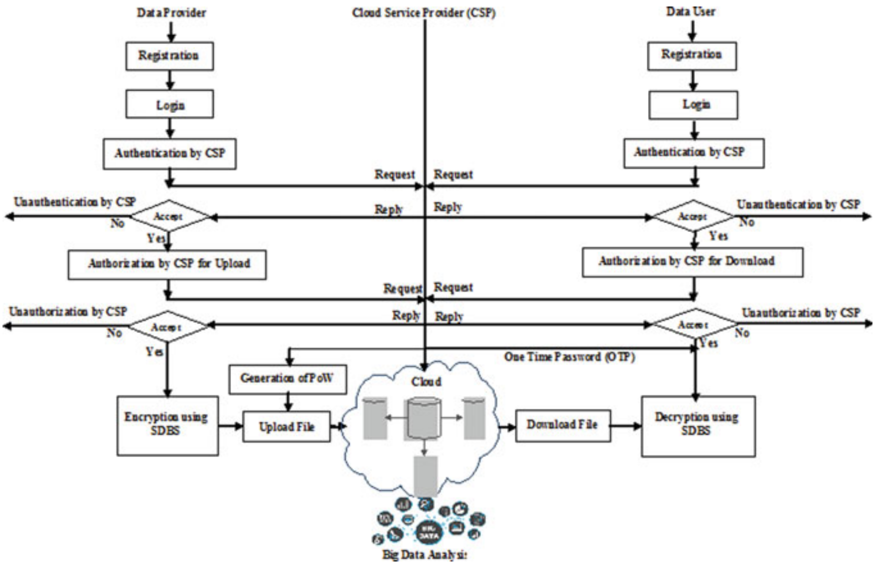



**Fig. 1** Enhanced security systems for big data cloud using SDBS

unauthenticated persons are irritating to login private to the organism. Completely security supply by the SDBS algorithm makes master key and session key to encrypt and decrypt the data [12].

## 5 Architecture Diagram for Enhanced Security System

The enhanced security architecture is illustrated in Fig. 1. A four-level data protection is ensured by applying various security services such as authentication, authorization, encryption, and decryption. Only the authenticated users are allowed to access the data from big data cloud.

The cloud users register their details in the registration form. As a first step, all the data providers and data users have to register their identity number (ID), username, password, role type, date of birth, mobile number, and e-mail ID into their systems. All these information will be stored in the big data cloud server [4].

## 6 Authentications of Big Data Cloud Users and Cloud Providers

Authentication is based on the personal details of the cloud user, and cloud service provider authenticates the cloud users. Authentication is the first level of security.

Data providers and data users are authenticated by cloud service provider. If the data provider wants to upload the data file, the data provider has to login to the system with username and password. Cloud service provider verifies and validates the username and password of the data provider with the database and accepts or rejects the request. This is called data provider authentication by cloud service provider or otherwise unauthentication by cloud service provider. In the same way, if the data user wants to download the data file from the big data cloud server, the user logins to the system with the username and password. Cloud service provider verifies and validates the username and password of the data user which is stored in a database based on accepts or rejects of the request of data user authentication by cloud service provider or otherwise unauthentication by CSP [4].

### *6.1 Authorization of Big Data Cloud Users and Cloud Providers*

Cloud service provider will authorize the cloud users by validating their identification accumulate in the record. Authenticated information provider sends a request to get the secret key to encrypt the data. Cloud service provider will authorize the data vendor by verifying their identification stored in the database. All the inventory users cannot upload the data. Solitary the authorized data supplier can upload the data [5].

### *6.2 Encryption of Uploaded Data Files and Decryption of Downloaded Data*

The data files are encrypted using Secure Dynamic Bit Standard algorithm and uploaded to the cloud server by the authorized cloud data providers. The data files are decrypted using SDBS algorithm and downloaded into the cloud server by the authorized cloud users. Only the legitimate data users can download and decrypt the file through one-time password received from the CSP [6].

## 7 Proposed Cryptographic Algorithm

Secure Dynamic Bit Standard is proposed for one-to-many encryption in which the identical plaintexts are encrypted in dissimilar behavior and we get dissimilar ciphertext. Encryption in Secure Dynamic Bit Standard is easy and secure compared to other encryption techniques. Still Secure Dynamic Bit Standard be related to Advanced Encryption Standard (AES) in its organization, this algorithm is simple since the complication of AES algorithm is detached. So, the method becomes easy and the instant convolution is also fewer. Secure Dynamic Bit Standard is secure because several individual of the principles are chosen arbitrarily at the instance of uploading, and the assailant gets enlightened concerning the amount of the encryption key and the operations. Intended for further security, two keys are used in this algorithm in which the master key is encrypted by the session key and solitary the authorized users preserve acquired the keys [5].

i. **Security**: In encryption, cryptography is demoralized. Symmetric key cryptography is demoralized for data encryption and asymmetric key cryptography is used for key creation.
ii. **Efficiency**: It uses equally the symmetric key and asymmetric key cryptography, and it is extremely protected and creates a competent result.
iii. **Data reliability**: The encrypted data can be decrypted by the authenticated users; the amendment or removal of the data is not potential.
iv. **Data performance**: Encryption or decryption of data is fast, and the uploading and downloading data performance are moreover best.

## 8 Results and Discussion

The proposed schemes are implemented on secure encryption and decryption. Figure 2 shows that the encryption time is required by the ABE, AES, hybridized ABE and AES and SDBS algorithms. The secure encrypted data is uploaded into the big data cloud by the data supplier. The data is downloaded by the data user and got decrypted to supply secure data storage and avoid data defeat. SDBS decrypted time consumes less when evaluated with ABE, AES, and hybridized ABE and AES algorithm.

Figure 3 illustrates the decryption time essential by the AES, ABE, and hybridized ABE and AES and SDBS algorithm. Encryption is used for enciphering the contented at sender end earlier than convey it over the system while decryption is used for deciphering the snarled empty satisfied at the receiver end.

Figure 4 clarifies evaluation of consequences of encryption instance agreed in millisecond with AES and hybridized ABE and AES algorithm of an integer of records uploaded on the cloud. Figure 4 illustrates to facilitate the encryption instance essential by the hybridized ABE and AES encryption algorithm necessitate significantly additional instance to encrypt evaluate with AES algorithm since of extra security. Consequently, it is to establish that the amount of records uploaded into the cloud is secure and protected.

Figure 5 illustrates the assessment of consequences of decryption time agreed in millisecond with AES and OTP through hybridized ABE and AES algorithm of an amount of client verification demand on the cloud.

Figure 5 illustrates to facilitate the decryption instance necessary by the OTP with hybridized ABE and AES encryption algorithm desires significantly more instance

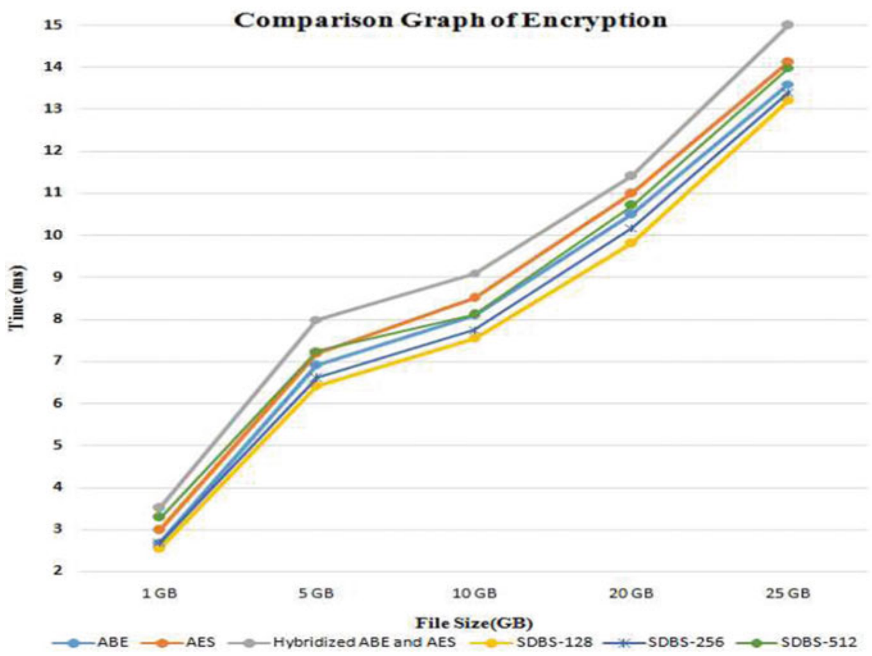



**Fig. 2** Evaluations of encrypted data

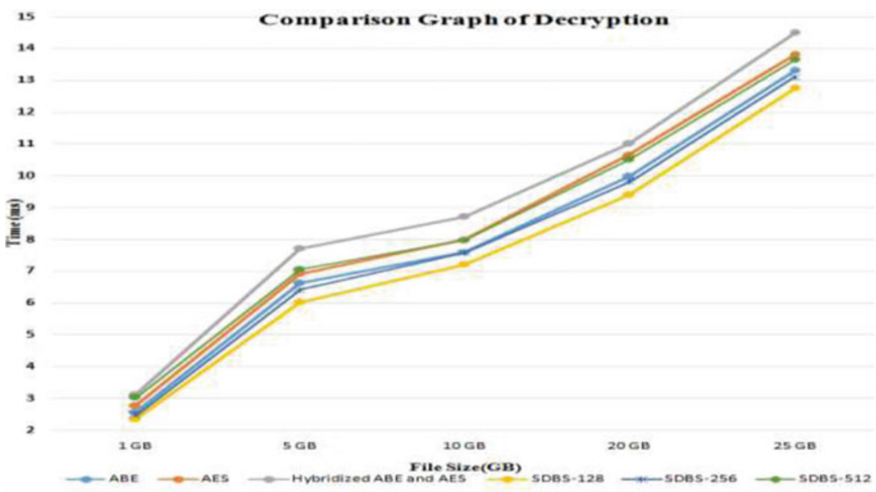



**Fig. 3** Evaluations of decrypted data

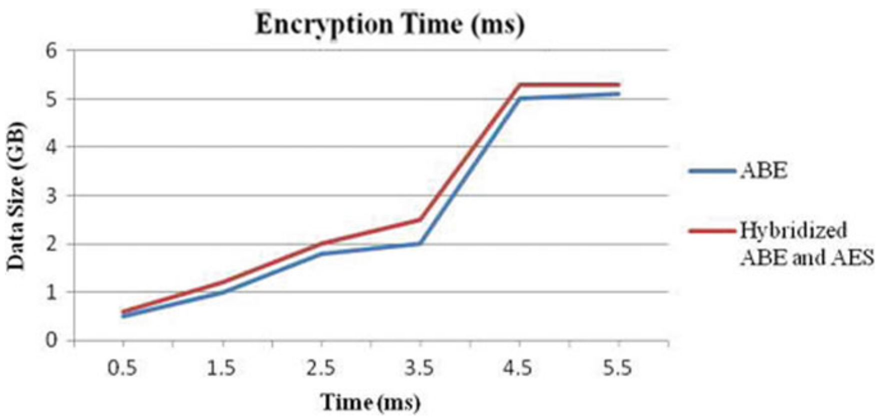



**Fig. 4** Comparisons of hybridized ABE and AES with ABE

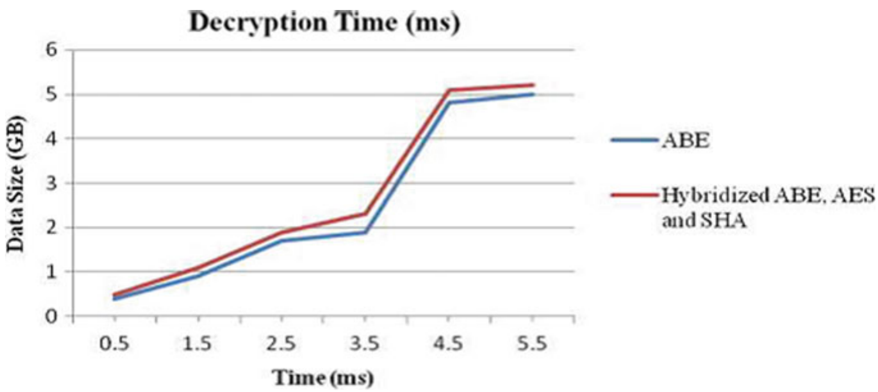



**Fig. 5** Comparisons of hybridized ABE, AES, and SHA with ABE

to decrypt evaluate with AES algorithm, and it depends on the client authentication request for downloading as of the cloud.

## *8.1 Big Data Analysis*

Figure 6 illustrates data supplier (DS) and data user (DU), encryption and decryption instance take using system 4 SDBS algorithm. Big data examination accumulate and evaluate five numbers of data supplier are DS1, DS2, DS3, DS4, and DS5. The total amount of bytes used by each data supplier or data user and data records is encrypted and decrypted by the use of SDBS algorithm.

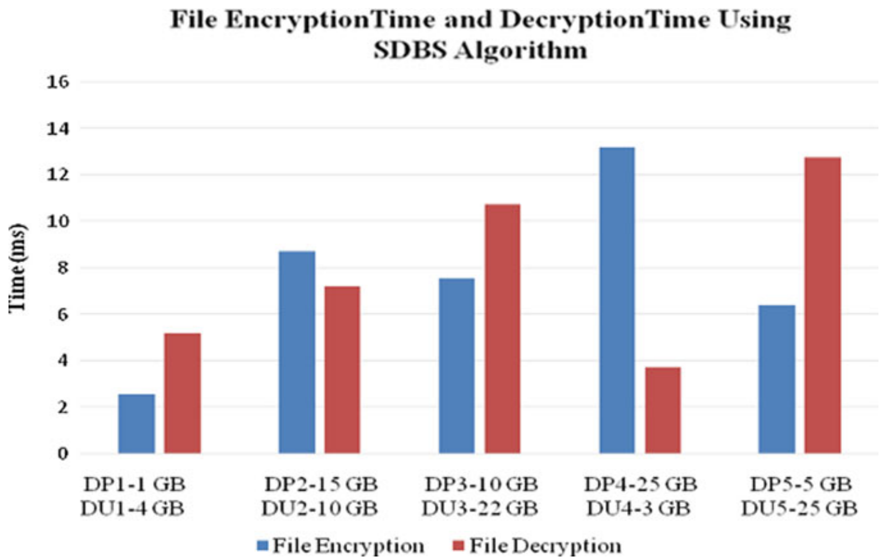



**Fig. 6** File encryption instance and decryption instance using SDBS 128-bit standard algorithm

DS1 uploaded completely 1 GB statistics file, DS2 uploaded completely 15 GB statistics file, DS3 uploaded completely 10 GB statistics file, DS4 uploaded completely 25 GB statistics file, and DP5 uploaded totally 5 GB statistics file. Their uploaded statistics file got encrypted time given as follows 2.56, 8.7, 7.55, 13.2, and 6.41 ms. Big data study that accumulates and compares five numbers of data users are DU1, DU2, DU3, DU4, and DU5. DU1 downloaded completely 4 GB statistics file, DU2 downloaded completely 10 GB statistics file, DU3 downloaded completely 22 GB statistics file, DU4 downloaded completely 3 GB statistics file, and DU5 downloaded completely 25 GB statistics file. Their downloaded statistics file got decrypted instance given as follows 5.2, 7.2, 10.75, 3.72, and 12.75 ms.

## *8.2 Encryption and Decryption Instance Investigation*

The statistics file encrypted and decrypted using ABE, AES, hybridized ABE and AES and SDBS creates enhanced privacy and security in cloud analyze. The advanced encryption and decryption instance is used to evaluate with ABE, AES, and SDBS algorithm. SDBS algorithm uses less encryption and decryption instance when evaluating with ABE, AES, and hybridized ABE and AES algorithms. ABE and AES algorithms obtain encryption and decryption instance in among hybridized ABE and AES and SDBS algorithms. Table 3 illustrates encryption instance and decryption instance show of scheme 1, scheme 2, scheme 3, and scheme 4 that are associated

**Table 3** Performance of ABE, AES, hybridized ABE and AES, SDBS 128-bit standard, SDBS 256-bit standard and SDBS 512-bit standard algorithms

| | Scheme 1—ABE | | Scheme 2—AES | | Scheme 3—Hybridized ABE and AES | | Scheme 4—SDBS 128-bits standard | | Scheme 4—SDBS 256-bits standard | | Scheme 4—SDBS 512-bits standard | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| File size (GB) | Encryption time (ms) | Decryption time (ms) | Encryption time (ms) | Decryption time (ms) | Encryption time (ms) | Decryption time (ms) | Encryption time (ms) | Decryption time (ms) | Encryption time (ms) | Decryption time (ms) | Encryption time (ms) | Decryption time (ms) |
| 1 | 2.69 | 2.56 | 2.99 | 2.76 | 3.5 | 3.1 | 2.56 | 2.35 | 2.69 | 2.47 | 3.29 | 3.03 |
| 2 | 3.41 | 3.27 | 3.71 | 3.57 | 3.91 | 3.77 | 3.28 | 3.17 | 3.37 | 3.25 | 3.88 | 3.63 |
| 3 | 4.42 | 4.23 | 4.52 | 4.43 | 4.82 | 4.73 | 3.98 | 3.72 | 4.1 | 3.9 | 4.64 | 4.45 |
| 4 | 5.45 | 5.36 | 5.65 | 5.5 | 5.95 | 5.86 | 5.32 | 5.2 | 5.49 | 5.39 | 5.91 | 5.78 |
| 5 | 6.9 | 6.62 | 7.2 | 6.92 | 8 | 7.7 | 6.41 | 6.02 | 6.61 | 6.42 | 7.25 | 7.05 |

to the different file size specified in gigabyte versus encryption, decryption instance given in milliseconds.

## 9 Conclusions

The data supplier has completed the data transfer to the big data cloud. The proposed scheme has high data veracity and data storage liberty exclusive of data defeat. Earlier than the information has been uploaded into the storage, a high protected algorithm called SDBS has been used. The safety investigation of different levels of security system, qualities and performance of big data cloud security scheme. Big data investigation information assists to recognize each data supplier and data user convention of file amount, encryption, decryption instance, and upload instance and download instance. This inspect has intended a confined scheme for big data cloud user. The statistics will be uploaded to the cloud solitary subsequent to authentic and authoritative by the cloud service provider. During a big data cloud condition, the data is accumulated in an encrypted scheme. The authenticated and authorized users can solitary decrypt the data subsequent to the login and one-time password compliance development. The register development needs a username, password, mail, address, phone number, and role category. For downloading the data from big data cloud, the data user desires to register and log in to make sure their authentication. Subsequently, the data user desires to send a download demand to the cloud service provider, for compliant the download demand the cloud service provider sends one-time password to data user mail id or mobile number. Afterward the data user wants to give one-time password if the provided one-time password value is correct; cloud service provider allows the data user to download the decrypted data from big data cloud. The innovative Secure Dynamic Bit Standard algorithm is planned for defensive the accumulate data. The algorithm is very complicated to assault by the hackers as the schemes occupy different levels of security. Formerly the authentication be unsuccessful it do not shift to subsequently step. Additionally the validation and endorsement schemes by the assist of the one-time password provide additional security. Though the concert of the scheme is good, transparency in time due to various steps is able to be concentrated advance by implementing modification such as declining the size of the key.

## References


2. Dai X, Dai H, Yang G et al (2019) An efficient and dynamic semantic-aware multi-keyword ranked search scheme over encrypted cloud data. IEEE Access 7:142855–142865
1. Ghanbarpour A, Naderi H (2020) An attribute-specific ranking method based on language models for keywordsearch over graphs. IEEE Trans Knowl Data En 32(1):12–25
3. Zhang X, Tang Y, Wang H et al (2019) Lattice-based proxy-oriented identity-based encryption with keyword search for cloud storage. Inform Sci 494:193–207

4. Yao J, Zheng Y, Wang C, Gui X (2018) Enabling search over encrypted cloud data with concealed searchpattern. IEEE Access
5. Nahar AK, Mongia K, Kumari S (2018) Cloud computing and cloud security. IJRAE&T 4(1):1–8
6. Uma G, Jayasimman L (2018) Enhanced convergent encryption key generation for secured data deduplication in cloud storage. J Phys Conf Ser 1142:012006
7. Fu Z, Xia L, Sun X, Liu AX, Xie G (2018) Semanticaware searching over encrypted data for cloud computing. IEEE Trans Inf Forensics Secur 99:1–1
8. Shrihari MR, Archana RA, Manjunath TN, Hegadi RS (2018) A review on different methods to protect big data sets. (12):4
9. Soumyashree, Nithyashree K (2017) A survey on security and privacy methods of bigdata in cloud computing environment. Int J Adv Res Ideas Innov Technol 3(6):1499–1504
10. Zhang Y, Chen X, Li J, Wong DS, Li H, You I (2017) Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing. Inform Sci 379:42–61
11. Yu Y, Au MH, Ateniese G, Huang X, Susilo W, Dai Y, Min G (2017) Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage. IEEE Trans IF&S 12(4):767–778
12. Chase J, Niyato D, Wang P, Chaisiri S, Ko R (2017) A scalable approach to joint cyber insurance and security-as-a-service provisioning in cloud computing. IEEE Trans Depend Sec Comput
13. Varsha V, Kousar H (2016) A survey on cloud computing its application and security issues. Int J Adv Eng Technol Manag Appl Sci 3(9):90–98
14. Casola V, De Benedictis A, Modic J, Rak M, Villano U (2016) Perservice security SLA: a new model for security management in clouds. In: IEEE 25th international conference on enabling technologies: infrastructure for collaborative enterprises, pp 83–88
15. Liu JK, Au MH, Susilo W, Liang K, Lu R, Srinivasan B (2015) Secure sharing and searching for real-time video data in mobile cloud. IEEE Netw 29(2):46–50
16. Hur J (2013) Improving security and efficiency in attribute-based data sharing. Knowledge and data engineering. IEEE Trans Knowl Data Eng 25(10):2271–2282

# Plant Leaf Disease Detection Using Image Processing

**M. Sahana, H. Reshma, R. Pavithra, and B. S. Kavya**

## 1 Introduction

Various diseases are observed during the growth of plants. The diseased plant or crop is identified by using its leaves. The different colored spots, blights, soft rots, and various patterns and marks on the leaf are very helpful in identifying the crop disease. The traditional scheme of identifying plant disease involves direct eye observation and recollecting certain set of diseases as per the season, climate, etc. The detection of crop disease is one of the ways to prevent the losses that occur in the crop yield. The quality and quantity of the agricultural outcome should be maintained.

The survey related to the plant diseases means the study of optically visible patterns on the plants. The crop disease detection and pesticide recommendation for a plant is very crucial. The sustainability of agriculture is very critical. It is very difficult to analyze and monitor the plant diseases manually. It requires tremendous amount of work to gain expertise in the plant disease detection. The system requires the excessive processing time.

Hence, image processing is the best tool for the prediction of plant diseases. The feature extraction is applied to the images during the data preprocessing. The CNN is used for classification and identification of the diseases. The estimation of pesticides and fertilizers is done by using TensorFlow technology.

The main objectives of this project are as follows:

- To identify the disease in plant leaf and estimate the fertilizers by using image processing.

M. Sahana (✉) · H. Reshma · R. Pavithra · B. S. Kavya
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore, India

B. S. Kavya
e-mail: kavya.bs@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_14

- Images are taken using a digital camera in the illumination of sunny days. The effect of illumination at the time of image capturing induces the illumination effect, but this is not considered at the time of image processing.
- Image processing tool box of OpenCV and convolution neural network is used for estimation of the fertilizers.

## 2 Related Work

The plant disease causes decrease in its growth and affects the photosynthesis. Photosynthesis activity generates the energy, which retains the growth and defense mechanism for survival [1].

About 58% smallholder farmers depend on agribusiness for food fodder and livelihood. The 80% of the emerging world depends on the agricultural products that are produced by the smallholder farmers. The 50% of major loss in the yield is due to the pest, pathogens, and diseases as described by Munzurul Islam [2].

The productivity is reducing gradually with various aspects. One of the major things is diseases on the crops that are not identified at the early stage. Bacterial and fungal diseases decrease the plant's growth very fast.

These types of diseases are studied by Bashish et al. [3]. They proposed the technique which identifies these types of diseases. The disease identification is done by image processing and color space transformation which generates the device discrete transformation.

The plant diseases that are detected at the early stages of growth would provide the maximum solution for the crop yield. Rastogi [4] has used ANN and K-means algorithm to classify and grade the diseases. The proposed automated system must identify the plant's disease and provide the recommendation of the correct pesticide.

Pesticides are applied to the rice crop for maintaining and controlling the diseases that affect the rice field [5]. Determining the correct pesticide for a particular disease is an important task [6] which gives the better solution.

A model which was built by Duthie [7] and Pennypacker et al. [8] is the best popular model for detecting the different type of diseases. The papers published by them use Weibull probability density function (PDF) and examine the outcomes of wetness period and the temperature of the plant. The proposed paper work describes the immediate identification of the plant diseases and affected section present in the leaf images of the grape plants by using support vector networks (SVM). This survey contains a unique work that will evaluate the percent of the infected plants.

Kaur [9] demonstrates the technical feasibility of machine learning approach to enable the user to detect the disease automatically. The disease diagnosis is done by image recognition. The disease status of 38 different classes containing 14 crop species and 26 diseases is done by achieving the accuracy over 99% by using deep convolution neural network, AlexNet, GoogLeNet, and stochastic gradient descent.

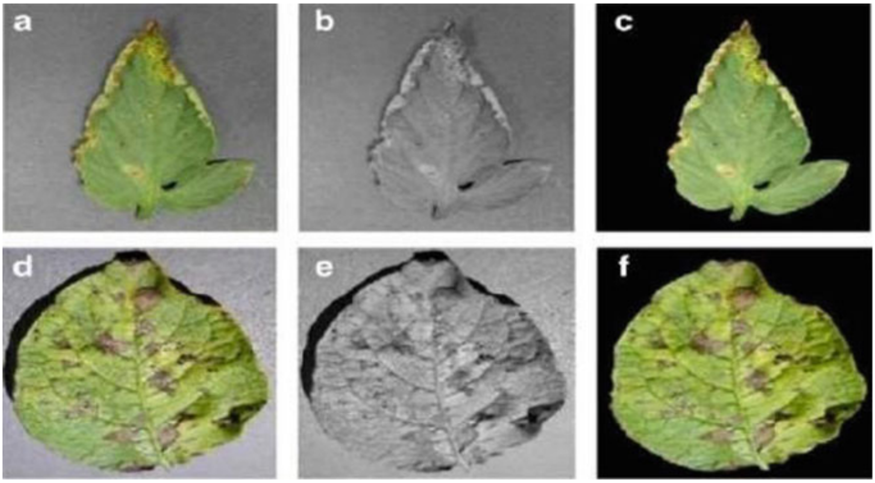



**Fig. 1** Different version of crop-disease pair image

# 3 Proposed System

## *3.1 Dataset Description*

The plant village dataset analyzes 54,306 images of crop leaves. They are divided into 38 class labels of crop-disease pair. Each class label is a crop-disease pair. The input images of plant leaves help the class label to predict the crop-disease pair. Three versions of datasets are used. First version of dataset has color images. Then, a grayscale view of the plant village dataset is used. In the final version, segmented analysis of the dataset is done. The background of the image provides the extra information to discover some of the inherent properties in the dataset.

For segmented dataset background information is removed, as shown in the below figures. Figure 1(a) Leaf 1 color, (b) Leaf 1 grayscale, (c) Leaf 1 segmented, (d) Leaf 2 color, (e) Leaf 2 Gy-scale, (f) Leaf 2 segmented.

## *3.2 Data Preprocessing*

The dataset obtained from the plant village Web site must be preprocessed before sending to the CNN model for training and testing process. The images in the dataset should be resized into 224 * 224 dimensions. The data preprocessing involves the following steps:

- Crop and adjust the division of the image.
- Flip the image horizontally.

- Adjust contrast, saturation, and hue.

As colored images are being used, the color conversion techniques are not required. The preprocessed images are directly sent to the CNN model for testing and training process. The images are taken from the plant village dataset for the training purpose. The GUI screen is used to fetch the image from the folder in the desktop during the testing process. The images are resized into 224 * 224 * 3 dimensions before passing it to the CNN for training.

### *3.3 Building a Convolution Neural Networks*

Different types of layers are used to build the convolution neural networks. They are:

- Convolution layer (CONV)
- Activation layer (ACT)
- Pooling layer (POOL)
- Fully connected layer (FC)
- Dropout layer (DO).

Considering all these layers, the convolution neural network is as follows:

$$\text{INPUT} \Rightarrow \text{CONV} \Rightarrow \text{RELU} \Rightarrow \text{FC} \Rightarrow \text{SOFTMAX}.$$

This model accepts the input, then moves on to the convolution layer, then to the activation layer, then a fully connected layer, and finally a softmax classifier. In softmax classifier, the output with the classification rate and algorithm possibility is obtained. The convolution and fully connected layers contain the parameters which are learned by the system through the training process.

The proposed system objective is to identify the disease in the crops grown at local region (INDIA). The next step is to make a pair of crop-disease and train the model again. After identifying the new pair of crop-disease, obtain the accuracy again. To achieve this, "crop disease net" architecture is used. "Crop disease net" architecture is specially designed to identify the disease from crop leaf image. As shown in Fig. 2, crop disease net architecture converts the crop image into $224 \times 224 \times 3$, where 3 is for a color image.

Crop disease net model consists of five convolution layers, three fully connected layers, and a softmax classifier. In crop disease net, the activation function ReLU is applied after every convolution and fully connected layer. Dropout is applied before the first and the second fully connected year. This network needs 1.1 billion computation units and has 62.3 million parameters. The convolution layers consume 95% of the computation and provide accurate results.

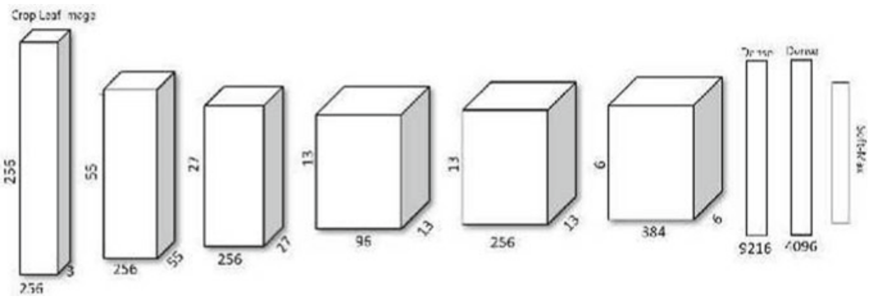



**Fig. 2** Crop disease net architecture

### *3.4 Architecture*

The pictorial representation of the complete project is shown in Fig. 3. The dataset is obtained from plant village Web site. The output of the project determines whether the leaf is healthy or unhealthy. In case of unhealthy leaf, disease name is determined and then the pesticide recommendation is displayed on screen.

## 4 Results

The proposed system identifies the plant diseases and also specifies the pesticide for the prevention of the disease. The feature extraction is applied to the images during the data preprocessing stage. The CNN is used for classification and identification of the diseases. The estimation of pesticides and fertilizers is done by using TensorFlow technology. Figures 4 and 5 represent the input image and prediction of the disease.

## 5 Conclusion

Plant protection in organic agriculture is not a simple process. The protection depends on how the plant is grown and preventing it from pests, pathogens, and weeds. This paper provides the deep learning models based on specific convolution neural network architecture. It is used for the identification of plant disease and also provides the recommendation of pesticides and fertilizers. Our experimental results show that how a CNN model can successfully recognize the different types of diseases in various plants and also suggest the suitable pesticides for the concerned disease. The plants with balanced nutrition and planted in good soil are able to resist pest/disease attacks successfully. The proposed system will contribute to agricultural research. In future, this work can be extended for a larger dataset by extracting more number of features.

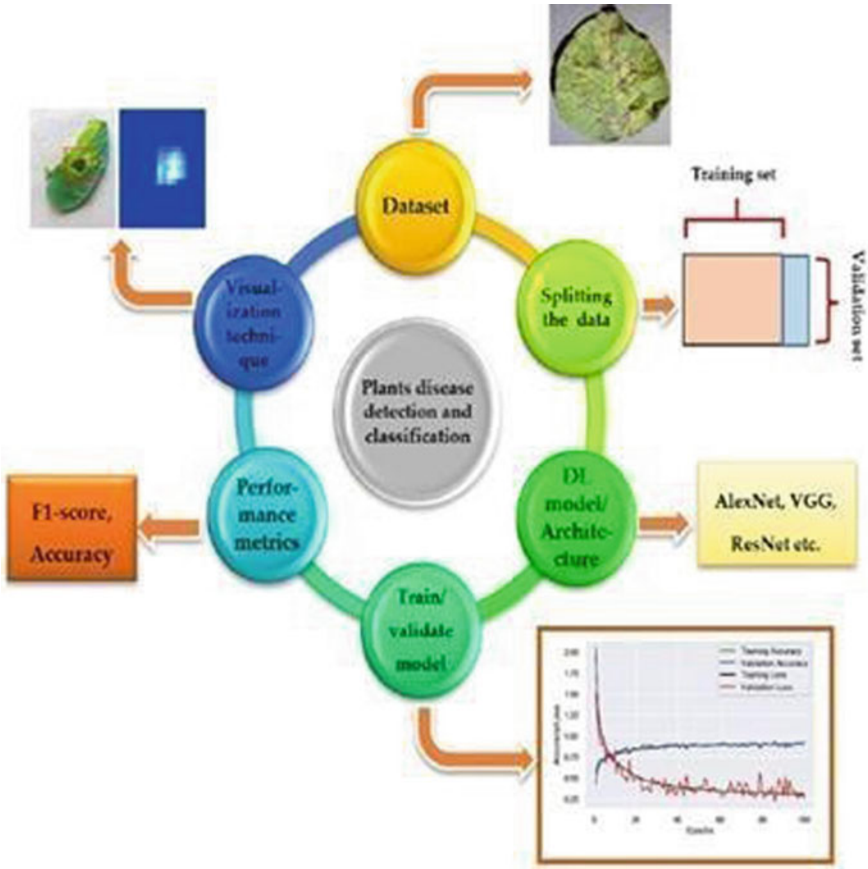



**Fig. 3** Functional flowchart

The Android application can be built for this work so that it can be used by the end user.


**Fig. 4** Input of image

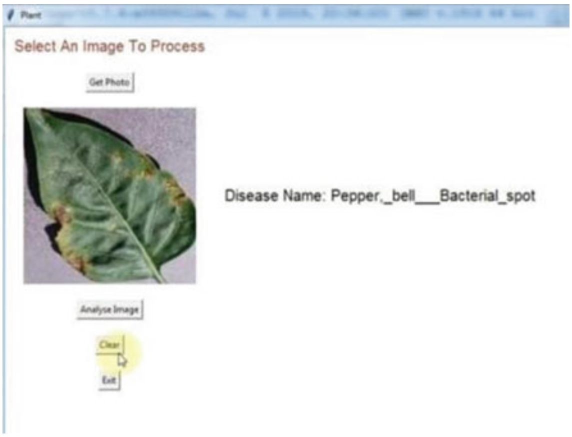



**Fig. 5** Prediction of the disease

## References

1. Management of leaf spot disease of trees and shrubs: Trees & Shrubs: Yard & Garden: Garden: University of Minnesota Extension, 4/22/2007
2. Islam M, Dinh A, Wahid K, Bhowmik P (2017) Detection of potato diseases using image segmentation and multiclass support vector machine. In: IEEE 30th Canadian conference on electrical and computer engineering (CCECE)
3. Al Bashish D, Braik M, Bani-Ahmad S (2010) A framework for detection and classification of plant and stem diseases. In: International conference on signal and image processing
4. Rastogi A, Arora R, Sharma S (2015) Leaf disease detection and Garding using computer vision technology and fuzzy logic. In: IEEE-IC on signal processing and integrated network
5. Phadikar S, Sil J (2008) Rice disease identification using pattern recognition. In: Proceedings of 11th international conference on computer and information technology

6. Neethu KS, Vijay Ganesh P (2017) Leaf disease detection and selection of pesticides using artificial neural network. Int Res J Eng Technol (IRJET)
7. Duthie JA (1997) Models of the response of foliar parasites to the combined effects of temperature and duration of wetness. Phytopathology 87–11:1088–1095
8. Pennypacker SP, Knoble HD, Antle CE, Madden LV (1980) A flexible model for studying plant disease progression. Phytopathology 70–3:232–235
9. Kaur R, Kang SS (2015) An enhancement in classifier support vector machine to improve plant disease detection. In: IEEE 3rd international conference on MOOCs, innovation and technology in education (MITE)

# Water Table Analysis Using Machine Learning

**S. Vijaya Shetty, Aishwarya Kulkarni, Shivangi Negi, Sumedha Raghu, C. V. Aravinda, and Guruprasad Hebbar**

## 1 Introduction

Groundwater is characterized as the water present in the sections between the soil pore spaces and within the cleavages of rock formations. The depth at which soil pore fractures and gaps in rock become saturated with water is determined as the water table [1]. In several arid and semi-arid areas, groundwater has emerged as an important source of water required for domestic, irrigation, urban, and industrial activities. India is a substantial consumer of groundwater in the world using 230 $km^3$ of groundwater per year, which is equivalent to the quarter of the global total [2]. Therefore, sustainable development of groundwater resources is essential for precise quantitative analysis, which is necessary for India due to its prevalent semi-arid and arid climate. Constant monitoring of groundwater levels is essential to prevent the misuse of groundwater resources that can usher to local water rationing, limiting in agricultural yields, wells going dried-up or generating unpredictable groundwater quality changes, variations in flow patterns of groundwater emerging in the inflow of meagre quality water and seawater intervention in coastal areas [3]. The water levels, if forecasted prior, might facilitate the executives to plan better, the groundwater usage. In this research, we focus on observation wells from various districts of Karnataka.

S. Vijaya Shetty (✉) · A. Kulkarni · S. Negi · S. Raghu
Nitte Meenakshi Institute of Technology, Bangalore, India
e-mail: vijayashetty.s@nmit.ac.in

C. V. Aravinda
NMAM Institute of Technology, Nitte, India
e-mail: aravinda.cv@nitte.edu.in

G. Hebbar
Dunn Solutions India Pvt. Ltd, Bengaluru, India
e-mail: guruprasad@dunnsolutions.com



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_15

Traditionally, process-based models are commonly used to perform groundwater simulation and predications, which rely on spatial data of the observed system dynamics. Though, they are not suitable in several arid and semi-arid areas as a consequence of insufficient data. Meanwhile, in data-driven modelling with machine learning methodology, our model attempts to establish a direct relationship among the inputs and outputs of the system without having any knowledge about the interior structure of the physical process [4]. The focus here is to use temporal data inputs (historic groundwater level, weather, and rainfall data) to learn the best approximation of the groundwater level values.

Recurrent neural networks (RNNs), a technique of deep learning, are a prominent choice for designing groundwater time series data due to their ability to retain a memory of previous network conditions, but they face challenges in acquiring long term dependencies within variables as weights associated with the network reaches to zero or turn exceedingly high while model training [5]. LSTM, a class of RNN can bypass these training problems by eradicating redundant information being passed to future model states while retaining a memory of relevant past events [5]. LSTM networks have lately incorporated to model the water table in an inland agricultural area of China on a monthly time step basis [6].

In our study, we are training models based on a set of related attributes to generate optimal predictions for the missing groundwater level values in our dataset using binary classification and later applying data-driven techniques to evaluate the performance of our models. Although data-driven approaches are increasingly used in surface water problems, there are hardly any studies related to groundwater in arid and semi-arid areas [4]. Therefore, the emphasis of this study is on the implementation of data-driven models with machine learning (i.e., RF AND GB) and deep learning (i.e., LSTM) and comparison of two ensemble methods (i.e., RF and GB) for forecasting groundwater levels in Karnataka, India.

## 2 Literature Survey

Various works are being implemented in the field of hydrology and groundwater study which also includes prediction analysis. The right form of data extraction is highly necessary for an accurate model [13], a complex web of factors that determines groundwater levels, which are: Rainfall, aquifers, precipitation levels, seasonal changes, groundwater storage patterns, water extraction. As highlighted in [7–9], Artificial Neural Networks (ANN) is ideal for forecasting based on the implementation of data from wells and shallow aquifers, respectively. Further, different algorithms are analysed [10]. Least Squares Support Vector Machine (LSSVM) was used for dynamic forecasting of wells in Mongolia. The paper [11] helped understand the wavelet integration with the support vector machine (SVM), which is a regression model that is being implemented to check the fluctuations in the water level. In [12], the focus is on quantitative estimates of groundwater temporally and spatially.

They have performed a study of groundwater level data in three districts of Maharashtra—Thane, Latur, and Sangli. Analysis of data of above 100 observation wells in each of these districts and developed seasonal models to represent the groundwater behaviour. Three different types of models were developed-periodic, polynomial, and rainfall models. In [13], a detailed survey was conducted concerning the prediction of groundwater levels of Ljubljana polje aquifer. Three different datasets from Ljubljana (Slovenia) and Skiathos (Greece): weather pump sensor data from Skiathos and, information. This paper processed the problem as a regression problem and hence implemented regression trees. They have highlighted the optimization when using ensembles by using random forests. Gradient boosting was also implemented. From the literature survey, neural networks were found to be widely used for predictive analysis. Hence for this Gradient Boosting and Random Forest are the suitable algorithms to be used for predicting the missing values and also compare the two algorithms to see which has better accuracy. LSTM (Long short-term memory) which is RNN (Recurrent neural network) model is considered for predicting future values.

## 3 System Architecture

The presented system as shown in Fig. 1, is developed using machine learning and deep learning application using which the dataset has been generated later analysed. The groundwater dataset is gathered from various aquifers and observational wells of several districts across Karnataka. The dataset which comprises of the multiple parameters of groundwater conditions is split into testing and training data. Firstly, cleaning of data is carried out by alternating the empty cells with the mean set of values. To compare the difference between a dataset with NULL values and that without one, the data is fit into the two algorithms that are Random forest and Gradient boosting. The loss functions MSE (Mean Square Error) is presented on the models. Also, the output obtained is plotted which highlights the most important parameter among all the existing parameters. Further, the LSTM (Long short-term memory network) is applied which is an RNN (Recurrent neural network) model to predict future water level values These results are intended to be given to the respective authority for necessary actions. Data pre-processing is the approach wherein data is gathered from the specific account using the XLS sheet. Since the gathered data is unstructured and not well defined, the pre-processing techniques are used which include normalization, data cleaning, dimension reduction. This pre-processed data is reviewed and analysed several times by considering all of its parameters before fitting it into the model. Next, the data is subjected to the prediction analyser using the algorithms that include gradient boosting, random forest, and LSTM namely. Finally, the outcome is compared to know how accurate the prediction is, also the result of the data cleaning and the replaced values is represented as graphs.

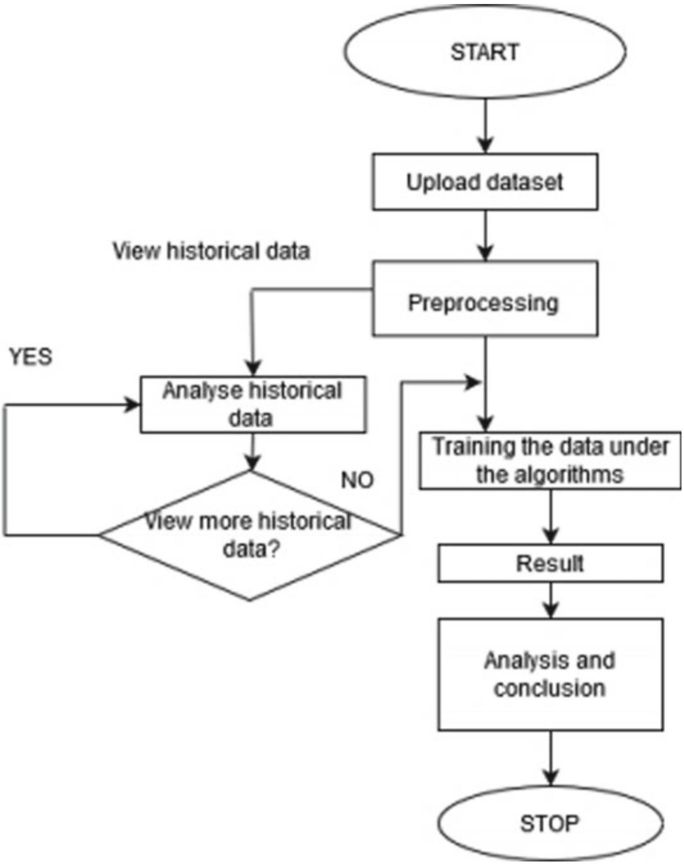



**Fig. 1** System architecture of groundwater level prediction

# 4 Dataset

Table 1 shows the dataset that has been used. The pre-processing is done on this dataset. The dataset consists of groundwater levels in the pre-monsoon season as well as the Post Rabi and Post Kharif crop seasons. It also consists of location data including the well-code, its district, state, site name, and site type. Table 2 represents the attributes in a given dataset.

## *4.1 Pre-Processing*

The data gathered is compiled from various government and water board sources. The dataset initially required a significant amount of pre-processing. The techniques

**Table 1** Dataset

| SITE_TYPE | WLCODE | YEAR_OBS | MONSOON | POMRB | POMKH | PREMON |
|---|---|---|---|---|---|---|
| BORE WELL | W05243 | 2018 | 19.63 | 18.11 | NA | NA |
| DUG WELL | W24336 | 2018 | NA | NA | 4.13 | 3.72 |
| BORE WELL | W05497 | 2018 | 23 | 24.57 | NA | NA |
| BORE WELL | W06424 | 2018 | NA | 59.32 | NA | NA |
| BORE WELL | W21201 | 2018 | NA | 59.3 | NA | NA |
| BORE WELL | W05727 | 2018 | NA | 21.06 | 37.5 | NA |
| BORE WELL | W05731 | 2018 | NA | 7.64 | 9.1 | NA |

**Table 2** Dataset description

| Attribute | Attribute description |
|---|---|
| STATE | Name of the state where the observational well is located |
| DISTRICT | Name of the district where the observational well is located |
| TEH_NAME | Name of the tehsil (administrative area) where the observational well is located |
| BLOCK_NAME | Name of the sub-regions in the district where the observational well is located |
| LAT | Latitude of the observational well location |
| LON | Longitude of the observational well location |
| SITE_NAME | Name of the site of the observational well |
| SITE_TYPE | Type of the site i.e. bore-well or dug-well |
| WLCODE | A specific unique number assigned to each observational well |
| YEAR_OBS | Year during which the observation has been recorded |
| MONSOON | Groundwater levels during the monsoon |
| POMRB | Groundwater levels post monsoon during the rabi crop season (October to November) |
| POMKH | Groundwater levels post monsoon during the kharif crop season (June to October) |
| PREMON | Groundwater levels before the monsoon season |

used on the groundwater dataset are data cleaning, data reduction, and checking feature importance. The dataset compiled contained a large amount of NaN (not a number) values, due to which data was not usable. We first replaced the NaN values with 0s and then used an imputer function to alternate the 0 s with the mean value of each column. Certain parameters like well code and site name were then dropped as it was not required for any of the methods employed in this project.

Then a feature importance was calculated for the four features YEAR_OBS, MONSOON, POMKH AND POMRB that contain groundwater level values. The plot shows POMKH (post monsoon kharif) (Fig. 2) has the highest importance. It is the top feature contributing to the model, while (year of observation) has the lowest

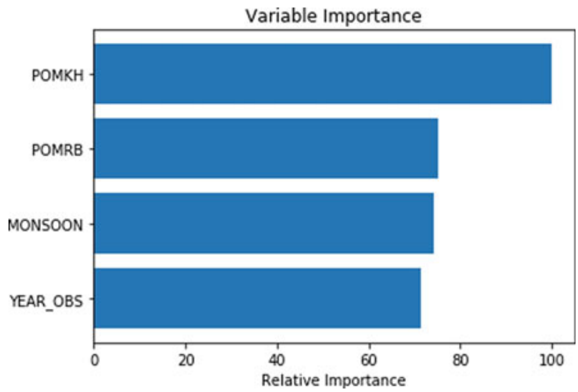



**Fig. 2** Feature importance after pre-processing

importance. When training trees one can compute the quantity, the feature contributes to decreasing the weighted impurity, which in case of regression trees is variance.

## 5 Algorithms

**Random Forest**

Random forest is one of the ensemble learning methodology that combines the concepts of classification and regression tasks with the help of multiple decision trees and a technique named bagging (Bootstrap Aggregation) with some additional degree of randomization.

Given a training set $Z = z_1, \ldots, z_n$ with responses $Y = y_1, \ldots, y_n$, bagging repeatedly ($M$ times) chooses a random sample with replacement and fits trees to generate new training sets:

For $m = 1, \ldots, M$:

- Sample, with replacement, n training examples from $Z$, $Y$ called $Z_m$, $Y_m$
- Train a classification or regression tree $f_m$ on $Z_m$, $Y_m$.

Bootstrap aggregation uses the following equation to predict unseen samples by averaging prediction from individual regression trees:

$$\hat{f} = \frac{1}{M}\sum_{m=1}^{M} f_m(z\prime) \quad (1)$$

where $\hat{f}$ represents the prediction for unseen samples.

An individual decision tree is considered as a weak predictor (low bias, high variance) but is fast enough to build. More trees give you a more robust model and prevent overfitting. When modelling, the data is resampled with replacement and for each sampling, a new classifier is trained. A new object type is classified based on new attributes and each tree votes a classification. The forest selects the classifications having the most votes of all the other trees in the forest.

In general, different classifiers overfit the data differently, and via voting, those disparities are averaged out.

### Gradient Boosting

Gradient boosting is a power technique for building predictive models. Gradient Boosting is about taking a model that by itself is a weak predictive model and combining that model with other models of the same type to produce a more accurate model. The idea is to compute a sequence of simple decisions trees, where each successive tree is built for the prediction residuals of the preceding tree.

In gradient boosting the weak model is called a weak learner. The term used when combining various machine learning models is an ensemble. The weak learner in XGBoost is a decision tree. The ensemble works in a forward stage-wise manner by adding classifiers one at a time such that the upcoming classifier is trained to improvise the previously trained ensemble, instituting a weak learner to improve the faults of the already present weak learners.

### LSTM

Long short-term memory (LSTM) units are units of a recurrent neural network (RNN) [5]. Recurrent Neural Networks suffer from short-term memory and may leave out important information from the beginning. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers the values over an arbitrary amount of time intervals and the three gates control and regulate the flow of information in and out of the cell. LSTM networks are used for classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series problem.

The forward pass of LSTM network is in the below equations:

$$\mathrm{fo}_t = \sigma_g(W_{\mathrm{fo}} x_t + C_{\mathrm{fo}} h_{t-1} + b_{\mathrm{fo}}) \tag{2}$$

$$\mathrm{ip}_t = \sigma_g\left(W_{\mathrm{ip}} x_t + C_{\mathrm{ip}} h_{t-1} + b_{\mathrm{ip}}\right) \tag{3}$$

$$\mathrm{op}_t = \sigma_g\left(W_{\mathrm{op}} x_t + C_{\mathrm{op}} h_{t-1} + b_{\mathrm{op}}\right) \tag{4}$$

Here $\mathrm{fo}_t$, $\mathrm{ip}_t$ and $\mathrm{op}_t$ can be described as forget gate, input gate and output gate, respectively. The matrices $W_q$ contains the weight of the input and $C_q$ contains recurrent connection and $\sigma_g$ is the sigmoid activation function. used in the LSTM network. The network output was calculated by stacking a fully connected layer on top of the LSTM cell. The product of the output layer is the forecast of the groundwater level for the coming season.

**Table 3** Confusion matrix from random forest regression

| | True positive | False negative |
|---|---|---|
| False negative | 2001 | 235 |
| True negative | 439 | 844 |

## 6 Methodology

The groundwater dataset initially contained 14 parameters (Table 2), and these parameters had a lot of NaN values. During pre-processing, we replaced the NULL values with 0. Using an imputer function, the 0 valued data with the mean value calculated for each column. The feature importance of the parameters containing numerical values was checked. We applied the random forest and gradient boosting algorithms on the dataset both before and after replacing the 0 valued data with the mean imputation. The accuracies of the two algorithms before and after the replacement of the null values were also calculated. We divided the dataset as a 70: 30 train-test split and applied the two techniques to determine accuracies, Mean squared errors, Confusion matrices, and classification reports.

For the random forest algorithm, we used the threshold value of 8.19, deliberated using the collective mean of all the groundwater values, to divide the data into 0 and 1 classes where any value below the threshold is classified as 0 and anything above the threshold is classified as 1. This division was created for the test and predicted classes to identify the data relationship between the true-positive, true-negative, false positive 4 and false negative, as shown in Table 3. In the test split that contained 3519 values, the following was observed. We then calculated the accuracy and MSE using the random forest regression algorithm. For gradient boosting algorithm, we made use of gradient boosting regressor function from the python libraries to calculate the accuracy and MSE of both before and after replacing the null values. These algorithms were also used to predict the missing data in the dataset. We then implemented an RNN network, LSTM i.e. Long Short Term Memory, to predict future groundwater levels, and to check the MSE values for training and test values, where we also check the accuracy of the predicted values compared to its test and train set values. By implementing the algorithms, we yield useful and necessary results.

## 7 Results

On extracting data from various sources, the data is cleaned and pre-processed, and is then fed to the random forest regressor model. To measure the effectiveness of the model the Confusion Matrix (Table 2) is provided. Let Truly Positive be TF, Truly Negative—TN, Falsely Positive—FP, Falsely Negative—FN, Actual results—(TP + FP) and Predicted results—(TP + FN); Based on the same it is found that the Recall (TP/Predicted results), Precision (TP/Actual results), f1-score (2 * (Precision * Recall)/Precision + Recall) and Accuracy (TP + TN/TP + TN + FN + FP) to

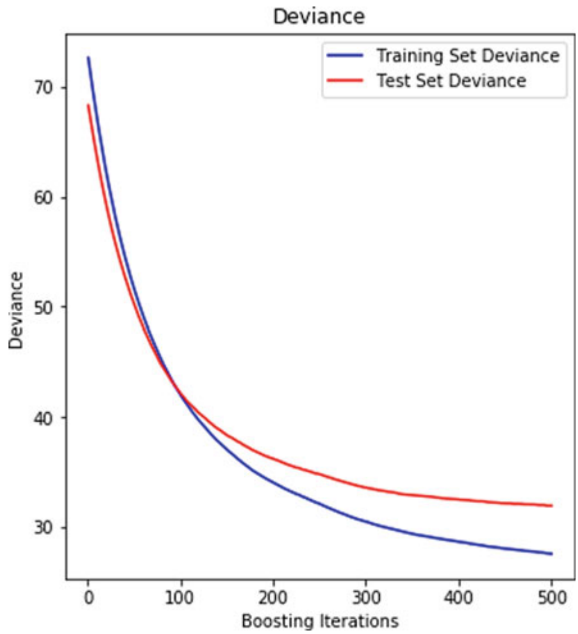
**Table 4** Comparison between random forest and gradient boosting

| Algorithm | MSE before replacement (out of 100) | MSE after replacement (out of 100) | Increase in accuracy (in %) |
|---|---|---|---|
| Random forest | 32.1665 | 20.3613 | 5 |
| Gradient boosting | 31.9301 | 19.5847 | 8 |

have increased from 0.81, 0.81, 0.81, 0.809 to 0.85, 0.85, 0.85 and 0.51 respectively after replacing the null values with mean values. Thus, we understand that all the measures of the confusion matrix have been in increased in their values which asserts the replacement.

On implementing gradient boosting we again have two conditions i.e. with the replacement of null values with mean and without replacement and we find that the MSE from 31.93 to 19.58 when replaced. In gradient boosting, it is found that in each iteration, we fit a base learner to the negative gradient of the loss function and later multiply the prediction with a constant and add it to the value from the previous iteration. From implementing the two algorithms, we observe that the MSE and accuracy using the gradient boosting yields the leading results for the dataset. Table 4 shows the Comparison between Random Forest Regression and Gradient Boosting. In the Gradient Boosting algorithm, we also checked for the training and test set deviance, where deviance is the goodness-of-fit statistic for a statistical model and we found that the deviance from the actual value is very minimal as seen in Fig. 3.

**Fig. 3** Data deviation as seen in gradient boosting



We have also implemented an RNN for prediction of future groundwater levels using the LSTM algorithm, which shows the improvement in the Root Mean Squared Error (RMSE). We can compare the predicted values with the existing one is in the dataset which is learnt by the machine learning model that we have used.

Out of every 100 data values, it is found that the RMSE score of the LSTM model for trained data compared to predicted values scored a 7.87, whereas for the tested data compared 5 to predicted values has scored a 13.36, which indicates how close the observed data points are to the model's predicted values. The overall performance of the models is found to be acceptable based on the high correlation efficiency.

Figure 4 is a scatterplot that depicts the groundwater level distribution in the 13 districts: Bangalore, Bangalore rural, Tumkur, Ramanagara, Kolar, Chikkaballapura, Mysore, Bijapur, Kodagu, Udupi, Bagalkot, Chikmagalur and Chitradurga. It is observed that regions such as Bijapur and Mysore receive lesser rainfall and therefore have lower groundwater levels than when compared to districts such as Kolar and Chikkaballapura. Thus it can be concluded that these trends are able to predict when a district is in critical condition with respect to groundwater levels.

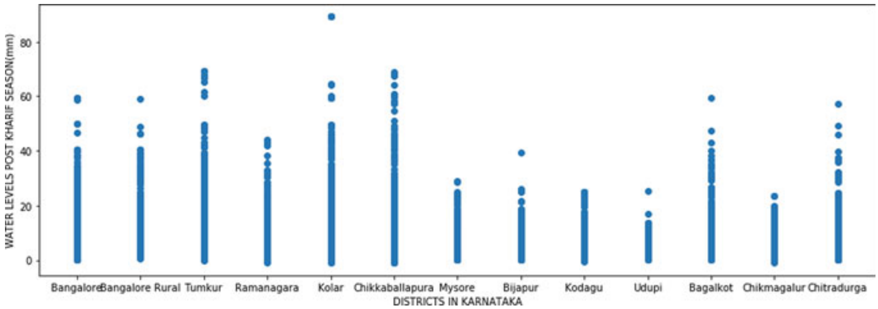
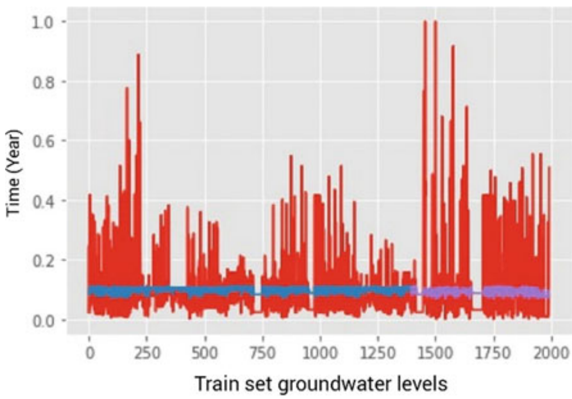



**Fig. 4** Groundwater level distribution in various districts

**Fig. 5** LSTM prediction using train set



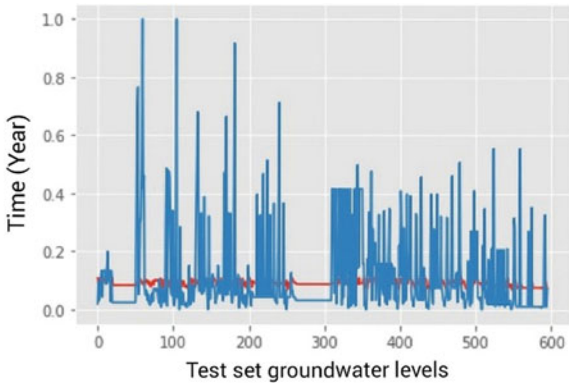



**Fig. 6** LSTM prediction using test set

Figure 5 is a graph comparing the train set values of the dataset (depicted in red) to the future values predicted by the LSTM network (depicted in blue) over the course of a year.

Similarly, Fig. 6 is a graph comparing the test set values of the dataset (depicted in red) to the future values predicted by the LSTM network (depicted in blue) over the course of a year. It is observed that the performance of the LSTM network is better with respect to test set values than with respect to the train set values. The overall performance of the models is found acceptable built on high correlation efficiency.

## 8 Conclusion

In this paper, we propose a groundwater level forecasting system. Three data-driven methodologies are tested based on various Machine Learning algorithms; namely, random forests, gradient boosting, and LSTM. The system is regarding using past groundwater levels data in the pre-monsoon and monsoon seasons as well as the Post Rabi and Post Kharif crop seasons to predict the groundwater levels for the missing values in the dataset, as well as for future usage.

Analysis of the results indicated that the designed gradient boosting model provided a good prediction of ground- water levels, with considerably good accuracy and lower value MSE. The methodology and findings demonstrated in this paper are useful to the research community of our nation involved in groundwater management and protection.

## References


1. Greenburg R (2005) The ocean moon: search for an alien biosphere. Springer Praxis Books
2. The World Bank, India groundwater: a valuable but diminishing resource, India-groundwater-critical-diminishing, 6 March 2012
3. Groundwater Governance. Thematic papers on groundwater. April 2016. PDF File
4. Huang M, Tian Y (2015) Prediction of groundwater level for sustainable water management in an arid basin using data-driven models. In: Proceedings of the 2015 international conference on sustainable energy and environmental engineering, pp 134–137. https://doi.org/10.2991/seee-15.2015.33
5. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
6. Zhang J, Zhu Y, Zhang X, Ye M, Yang J (2018) Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas. J Hydrol 561. https://doi.org/10.1016/j.jhydrol.2018.04.065
7. Mao X, Shang S, Liu X (2002) Groundwater level predictions using artificial neural networks. Tsinghua Sci Technol 7(6):574–579
8. Nayak PC, Rao YRS, Sudheer KP (2006) Groundwater level forecasting in a shallow aquifer using artificial neural network approach. Water Resour Manag 20:77–90
9. Yang Z, Lu W, Long Y, Li P (2008) Application of back-propagation artificial neural network models for prediction of groundwater levels: case study in Western Jilin Province, China. In: 2008 2nd international conference on bioinformatics and biomedical engineering, Shanghai, pp 3203–3206. https://doi.org/10.1109/ICBBE.2008.1130
10. Liu J, Chang J, Zhang W (2009) Groundwater level dynamic prediction based on chaos optimization and support vector machine. In: 2009 third international conference on genetic and evolutionary computing, Guilin, pp 39–43. https://doi.org/10.1109/WGEC.2009.25
11. Suryanarayana Ch, Sudheer Ch, Mahammood V, Panigrahi BK (2014) An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India, Neurocomputing 145:324–335. ISSN 0925-2312
12. Lalit K (2012) Temporal models for groundwater level prediction in regions of Maharashtra, June 2012. PDF File
13. Kenda K, Senozetnik M, Klemen K, Mladenić D (2018) Groundwater modeling with machine learning techniques: Ljubljana Polje Aquifer. Proceedings 2. https://doi.org/10.3390/proceedings2110697

# A Custom Classifier to Detect Spambots on CRESCI-2017 Dataset

**Karthikayini Thavasimani and N. K. Srinath**

## 1 Introduction

Online social networks such as Twitter, Facebook, LinkedIn, and Quaker play a vital role in our lives as one of the primary modes of communication. They are a great community platform for users to connect, discuss, and share their opinions. Nevertheless, it's being exploited by the social spam bots due to the popularity, ever-growing data, and rich API in the social networks [1]. Ferrara [2] describes a view of malicious activities played by bots and the current bot detection methods. Florian et al. explained [3], how bots act like humans and discussed various harms such as psychological, legal, economic, social, and democratic that bots create. Different strategies can be applied to prevent and protect bots from the above. Content analysis techniques like natural language processing and machine learning techniques help in detecting the bots that are grossly offensive, obscene, indecent, threatening, disclosing private information, false allegations and spreading misinformation. Behavior analysis techniques such as analyzing network traffic between source and destination IPs, rate of communication frequency between IPs, Social network/graph analysis, Botnet, or malware detection can help in preventing economic and democratic harms. Because of depth, the computing fraternity has been developing advanced techniques to detect social bots accurately. These techniques [4] are classified into three categories, structure-based bot detection, crowd-sourcing-based bot detection and machine learning-based bot detection. In 2017, the structure-based techniques were able to identify the bots with 68.2% accuracy and machine learning-based techniques with 86% accuracy. These techniques face several limitations such as big data processing time, high running time cost, decrease in false positive (FP) and false

K. Thavasimani (✉) · N. K. Srinath
RV College of Engineering, Bangalore, India

N. K. Srinath
e-mail: srinathnk@rvce.edu.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_16

negative (FN) rates, inefficient handling of AI-powered bots, Unknown source of Sybil accounts and so on.

This paper [5] is about ongoing US congress investigation accusing Russia of using trolls (malicious accounts) and BOTS (automated accounts) for spreading misinformation, politically biased information to influence the public. A Twitter dataset with over 43 million election trolls was collected from September 2016 to November 2016, with about 5.1 million distinct users. Retweet network and Label propagation algorithms are applied to classify liberal and conservative users with precision and recall of over 90%. Significantly, Conservatives are the ones who retweeted the Russian trolls over 36 times than the liberals. However, the label propagation algorithm is not detailed enough to understand the classification. Botometer, a machine learning framework that extracts and analyzes various features such as network structure, spanning content, text, and sentiments, is used to detect social bots. They estimated that 4.9 and 6.2% of liberal and conservative users were bots. Geospatial analysis is used to explore the location of users and identified that most of the tweets originated from southern states.

In reference [6], the author targets twitter data for identifying bots by applying Bot Classifier and showed that the algorithm identifies twitter bots with 83% accuracy than the traditional naive Bayes classifier. The author focused on twitter bot detection as a supervised classification problem and used various existing machine learning algorithms after extensive data pre-processing and feature extraction. A total of 62 features (from DARPA twitter bot challenge) were extracted based on user information, tweets, and periodicity. The extracted features were then fed into machine learning algorithms to see comparative effects. Gradient boosted trees outperformed others with 82% F1 score and 86% accuracy [7]. Cai et al. proposed a framework BeDM [8] (Behavior enhanced Deep Model) to identify a unique pattern to classify humans and bots. BeDM outperformed existing Bot detection analysis frameworks such as stweeler, Boosting, BoostOR with 87% accuracy.

In the paper [9], a method to identify spam bots based on empirical analysis and JavaScript testing framework was proposed and were able to identify 11% of spambot accounts. Two-class Boosted Decision tree classification model was proposed and repeatedly trained with website-specific hypotheses to detect bots to achieve an accuracy of 83% [10]. After reviewing the literature surveys to the best of our knowledge, we could identify that most of the existing algorithms were able to classify the bots with 83–87% of accuracy. These might be related to issues such as imbalanced dataset, pre-processing, feature selection or the learning problem of the model. This paper [16] uses various intrusion detection approaches such as statistical, artificial neural networks, support vector machines, data mining, rule-based systems and fuzzy logic to detect the threats in the networking sites. In addition to that, it explains the core differences between machine learning and deep learning techniques.

Intrusion detection systems play a crucial role in preventing various attacks on networking sites and can eventually prevent spam bots in social networking sites. The author [17] explained multiple reasons on why spambots are created in social networking sites. They are as follows.

1. Publish phishing links or malware sites
2. Aggressive following behavior (without mass follow up and mass attention)
3. Persecution of answers or functions to send spam messages to the user
4. Create duplicated accounts (manual or automated tools using)
5. Trolling trending topics
6. Sending replicated updates
7. Publish links with irrelevant tweets.

## 2 Open Research Problems

There are a lot of challenging issues in detecting spambots on social networks such as,

(1) Big data is dynamically growing in the social network. In the case of structure-based detection schemes [4], the techniques utilize high running time cost even with static data. Thus, a need for a computationally more efficient real-time graph algorithm is a rising good research challenge to handle big data.
(2) Most of the existing machine learning techniques [4] produce high false positive rates (FP) with comparatively low accuracy. For example, SybilFrame produces 4.2% false positives (FP), with a classification accuracy of 95.4%. Henceforth, a new learning approach can be employed to reduce FP (false positive) and FN (false negative rates).
(3) Nowadays bots are highly equipped with AI power. In the case of crowd-sourcing-based detection, protecting user privacy [4] is the biggest challenge.
(4) Additional features can be employed in machine learning technique to trace the bots accurately as bots are evolving with human-like behavior. Deep learning techniques [4] will be more appropriate to distinguish human from bot.
(5) Newly evolving bots such as scraper bots (price, ticket, content), comment spam, credential abuse, DDoS attack bots are sprouting in the web services, social media platform, apps, etc. can be another research challenge.

## 3 Data Collection

In our proposed work, we used the dataset from CRESCI-2017 [13, 14] that contains a combination of genuine user accounts and spambots such as social and traditional. Every group name contains a user and tweet file. The user file contains the information of the users. The tweet file contains a collection of tweets posted by individual users.

A total of 7931 twitter accounts and its user's corresponding tweets of 3,604,238 is analyzed. [14] The accounts are classified using digital DNA and manual analysis. A combination of LCS (Longest Common Substring) curve, available user variables and various twitter entities such as hashtag, media, URL, user mentions, and retweets are used for detecting the spambots. MCC (Matthews Correlation Coefficient) uses all the elements of the confusion matrix, and the reported values are higher denoting the accuracy of the labeled dataset. For testing the existing machine learning algorithms and the impact of the custom metrics on scalar data, a labeled dataset is crucial in determining the number of false positives, false negatives, true positives and true negatives. So, we used the CRESCI 2017 dataset. However, CRESCI-2019 dataset is not used in our analysis due to lack of tweet-level features.

## 4 Data Pre-processing

The selected dataset contains sparse data with lots of missing values. The data is a mix of text and numbers. So, we used categorial feature selection methods. In step 1, the missing values are filled with default values such as 0 for numbers and NA for strings. The dataset contains a mix of user and tweet specific features.

**Step 1**
The user file is a categorical file with 40 features. Crawled-at and timestamp are ignored since they don't impact the bot classification. Then, a new feature called "target value" is added with values 0 or 1 based on the group name mentioned in Table1, i.e., the social and traditional spam bots files are filled with "1" indicating bot accounts whereas genuine accounts and fake followers' target value is filled with "0" indicating human accounts. Finally, the user file consists of 39 features with missing data. Multiple approaches are available to handle missing values such as removing missing values or imputing them with sensible values and so on. We replaced the missing values with 0 for numerical data and NA for non-numerical data using fillna() function of the Pandas library. The pre-processed data of all groups are merged into a single file named as USER_INPUT.

**Table 1** Descriptive statistics on each dataset

| Group Name | Description | Accounts | Tweets | Year |
|---|---|---|---|---|
| Genuine accounts | Verified accounts that are human-operated | 3474 | 8,377,522 | 2011 |
| Social spambots #1 | Retweeters of an Italian political candidate | 991 | 1,610,176 | 2012 |
| Social spambots #2 | Spammers of paid apps for mobile devices | 3457 | 428,542 | 2014 |
| Social spambots #3 | Spammers of products on sale at Amazon.com | 464 | 1,418,626 | 2011 |
| Traditional spambots #1 | Training set of spammers used by C. Yang, R. Harkreader, and G. Gu | 1000 | 145,094 | 2009 |
| | | | 7931 | 3,602,438 |

**Step 2**

The tweet file is also a categorical file with 25 features. A total of 3,604,238 tweets are pre-processed. Due to memory limitations to read complete data, we used batch-processing. Larger files are processed in batches to improve processing speed. As a result, we were able to pre-process the data using a regular desktop with 12 GB RAM. Numerical features such as possibly_sensitive, retweet_count, reply_count, favorite_count, num_hashtags, num_urls and num_mentions are used for analysis. The remaining features are ignored since they don't impact the bot classification. A single user is associated with a collection of tweets in tweet file. So, for every user, the selected numerical features are imputed with the mean value. The output contains the mean values of the selected features for individual users and exported into a new file with name, TWEET_INPUT (Tables 2 and 3).

**Table 2** User level features

| id | name | screen_name |
|---|---|---|
| statuses_count | followers_count | friends_count |
| favourites_count | listed_count | url |
| lang | time_zone | location |
| default_profile | default_profile_image | geo_enabled |
| profile_image_url | profile_banner_url | profile_use_background_image |
| profile_background_image_url_https | profile_text_color | profile_image_url_https |
| profile_sidebar_border_color | profile_background_tile | profile_sidebar_fill_color |
| profile_background_image_url | profile_background_color | profile_link_color |
| utc_offset | is_translator | follow_request_sent |
| protected | verified | notifications |
| description | contributors_enabled | following |
| created_at | timestamp | crawled_at |
| updated | Bot_(added) | |

**Table 3** Tweet-level features

| id | text | source |
|---|---|---|
| user_id | truncated | in_reply_to_status_id |
| in_reply_to_user_id | in_reply_to_screen_name | retweeted_status_id |
| geo | place | contributors |
| retweet_count | reply_count | favorite_count |
| favorited | retweeted | possibly_sensitive |
| num_hashtags | num_urls | num_mentions |
| created_at | timestamp | crawled_at |
| updated | | |

**Step 3**

As a result of step1 and step 2, we have equal number of rows in both USER_INPUT & TWEET_INPUT with 46 features. So, they are merged to a single file called as "MASTER_INPUT". The final step is replacing the string values with 1 if not-null, otherwise 0. The final feature set after data pre-processing are id, name, screen_name, statuses_count, followers_count, friends_count, favorites_count, listed_count, url, lang, time_zone, location, default_profile, default_profile_image, geo_enabled, profile_image_url, profile_banner_url, profile_use_background_image, profile_background_image_url_https, profile_text_color, profile_image_url_https, profile_sidebar_border_color, profile_background_tile, profile_sidebar_fill_color, profile_background_image_url, profile_background_color, profile_link_color, utc_offset, is_translator, follow_request_sent, protected, verified, notifications, description, contributors_enabled, following, created_at, updated, possibly_sensitive, retweet_count, reply_count, favorite_count, num_hashtags, num_urls, num_mentions and bot.

## 5 Feature Selection

Feature selection is the process of selecting crucial features or attributes that are most relevant to the target variable. As mentioned above, MASTER_INPUT file is a categorical dataset with 46 features. Thus, the most popularly used categorical selection techniques such as Chi-squared and Mutual Information are used.

Chi-squared technique listed the below features as crucial for prediction. 24 out of 40 features are listed as important for prediction (Table 4).

Similarly, mutual information listed the below features as important with correlation threshold above 0.05 (Table 5).

Of 40, only 7 features are marked as crucial for prediction. So, we used chi-square feature selection in our proposed classification model.

**Table 4** Chi-square feature list

| name | statuses_count | followers_count |
|---|---|---|
| favourites_count | listed_count | url |
| lang | time_zone | location |
| default_profile | geo_enabled | profile_banner_url |
| profile_background_image_url_https | profile_text_color | profile_image_url_https |
| profile_sidebar_border_color | profile_sidebar_fill_color | "profile_background_image_url" |
| profile_background_color | profile_link_color | "utc_offset" |
| protected | verified | description |

**Table 5** Mutual information feature list

| | | |
|---|---|---|
| statuses_count | favourites_count | geo_enabled |
| protected | num_mentions | num_urls |
| num_hashtags | | |

## 6 Proposed Method

We proposed a new bot classifier algorithm based on Euclidean squared distance. The reason for using the distance metric is that it works well with sparse data and numerical values. The pre-processed dataset contains a numerical representation of the selected features. Since the values are numerical, determining the distance between the trained and test dataset will help us in predicting the target output efficiently. Hence, we propose that the Euclidean distance will be the appropriate metric to determine the target output (bot or not).

The Euclidean distance function measures the distance between the data points. In mathematical terms, the distance between the point $X$ and $Y$ is calculated as follows

$$D = \sqrt{\sum_{i=1}^{n} (xi - yi)^2}$$

where $d$ is the output.

For our implementation, we used Scikit learn [15] and Python. The Chi-Square feature selection is ingested into the code as the input. The proposed method has three modules, 1. Core 2. Estimator 3. BOT classifier. The core is the main module that calls the estimator and bot classifier algorithm for classification. The data is split into training and test dataset using random function. 80% of the data is used for training the algorithm, and the rest 20% is used for testing the prediction accuracy. The estimator checks the data for null/empty vales and updates them with default values (Numeric—0; Non-numeric: NA) in case of missing data. It helps us in avoiding overfitting and underfitting the data. First, the BOT classifier algorithm is trained using the training dataset. Then the test data is passed to verify the classification. Metrics such as precision, recall, and accuracy [20] are used to determine the efficiency of the algorithm. The BOT classifier results are compared with existing machine learning algorithms such as decision tree, $k$-nearest neighbor, linear discriminant analysis, Gaussian Naive Bayes and support vector machine.

Module 1: Core

Read the input csv file (preprocessed)
Select the features
Store the dataset with selected features to train and predict the output
Split the dataset into training and test data using random method
Call the Estimator
If the data is fit
Call the KNIClassifier
Train the algorithm using training dataset
Predict the training dataset results
Predict the test dataset results
Determine the accuracy of the predicted results

Module 2: Estimator Algorithm

Input = Preprocessed Dataset
For each entry in the input
Read the input
Check the values for null or empty
Fill the null or empty values with default values (neighboring values in case of non-numerical)
End For
Determine the fit status of the input
Return the status

Module 3: BOT Classifier Algorithm

Input = Estimator and Classifier mixer
Feed the dataset
For each Row in the Dataset
Determine the Euclidean distance between the data points (training and test data)
Based on the distance, select the nearest data point result
Add the results to the bot column in the dataset row
End For
Return resultant output dataset

## 7 Experimental Results

Evaluation metrics such as True Positive Rate (TPR), True Negative Rate (TNR), False Positive Rate (FPR), False Negative Rate (FNR), Positive Prediction Value (PPV), Negative predictive value (NPV), Positive likelihood (LR+), Negative likelihood (LR−), Diagnostic odds ratio (DOR), Youden's index (YI), Matthews correlation coefficient (MCC), Jaccard, Accuracy, Precision, recall and Error rate are used [20]. The proposed algorithm outperforms existing machine learning classifiers such as decision tree, K-NN, Linear discriminant analysis, Gaussian Naïve Bayes, Support vector machine in terms of test accuracy.

## 8 Conclusion

The rising problem of the social network is exploited by the malicious activities performed by bad bots to disseminate the information, influencing the public, fake advertisements about the products, psychological threats to users, politically biased information, stock exchange frauds and so on. This eventually rising another issue among the data scientists or researchers to yield the quality data for their analysis. In this regard, we discussed the existing techniques to detect bots and open research problems and challenges in this paper.

In our proposed work, we used the CRESCI-2017 dataset published by Indiana University that is a collection of genuine, traditional and social spam bot accounts. The data is pre-processed, features are selected and then fed into the Bot classifier algorithm to determine bots and non-bots. The algorithm comprises of three modules such as core, estimator, and Bot classifier. Estimator ensures that the data is fit, and Bot classifier predicts the output using Euclidean squared distance. Based on the distance, the nearest data point is selected as a result. Various classification metrics are used to analyze the performance of the algorithm with existing machine learning classifiers. The proposed algorithm outperforms the existing algorithms in terms of accuracy, error rate, precision, recall, and other metrics as listed in Table 6. Some of the limitations of the existing distance-based machine learning classifiers are lack of implementation and support for categorical data. As future work, we are working on extending this algorithm to support categorical data with deep learning on behavioral dataset such as network data.

**Table 6** Classification metrics

| Machine learning classifiers | | | | | | |
|---|---|---|---|---|---|---|
| Metrics | Decision Tree | k nearest neighbour | Linear discriminant analysis Gaussian Naive Bayes | | Support vector machine | Bo classifier |
| Accuracy | 81.50 | 94.00 | 92.00 | 89.00 | 91.00 | 96.13 |
| Precision | 0.89 | 0.94 | 0.92 | 0.89 | 0.92 | 0.96 |
| Recall | 0.82 | 0.94 | 0.92 | 0.89 | 0.91 | 0.96 |
| TPR | 0.56 | 0.89 | 0.83 | 0.87 | 0.95 | 0.92 |
| TNR | 0.98 | 0.95 | 0.95 | 0.89 | 0.90 | 0.97 |
| FPR | 0.01 | 0.04 | 0.04 | 0.10 | 0.09 | 0.02 |
| FNR | 0.43 | 0.10 | 0.16 | 0.12 | 0.04 | 0.07 |
| PPV | 0.96 | 0.86 | 0.84 | 0.63 | 0.66 | 0.90 |
| NPV | 0.77 | 0.96 | 0.95 | 0.97 | 0.98 | 0.97 |
| FDR | 0.03 | 0.13 | 0.15 | 0.36 | 0.33 | 0.09 |
| LR+ | 2.24 | 8.85 | 5.83 | 7.30 | 18.49 | 13.73 |

(continued)

**Table 6** (continued)

| Machine learning classifiers | | | | | | |
|---|---|---|---|---|---|---|
| Metrics | Decision Tree | k nearest neighbour | Linear discriminant analysis Gaussian Naive Bayes | | Support vector machine | Bo classifier |
| LR- | 0.44 | 0.11 | 0.17 | 0.13 | 0.05 | 0.07 |
| DOR | 5.05 | 78.43 | 34.00 | 53.40 | 341.89 | 188.73 |
| YI | 0.54 | 0.85 | 0.78 | 0.77 | 0.85 | 0.90 |
| MCC | 0.63 | 0.83 | 0.79 | 0.68 | 0.74 | 0.89 |
| Jaccard | 0.55 | 0.77 | 0.72 | 0.58 | 0.64 | 0.84 |
| Error rate | 0.18 | 0.05 | 0.07 | 0.10 | 0.08 | 0.03 |

## References

1. Dehade K, Bagade M (2015) A review on detecting automation on Twitter accounts. Eur J Adv Eng Technol 2:69–72
2. Ferrara E (2016) The rise of social bots. Commun ACM 59(7):96–104. https://doi.org/10.1145/2818717
3. Daniel F, Cappiello C, Benatallah B (2018) Bots acting like humans: Understanding and preventing harm. IEEE
4. Karataş A, Şahin S (2017) A review on social bot detection techniques and research directions. https://www.researchgate.net/publication/322853694
5. Badawy A, Ferrara E, Lerman K (2018) Analysing the digital traces of political manipulation: the 2016 Russian interference Twitter campaign. In: IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)
6. Tiwari PK, Velayutham T (2017) Bot identification: helping analysts for right data in Twitter. IEEE
7. Kantepe M, Ganiz MC (2017) Preprocessing framework for Twitter bot detection. In: IEEE 2nd international conference on computer science and engineering
8. Cai C, Zeng D (2017) Behavior enhanced deep bot detection in social media. In: IEEE international conference on intelligence and security informatics (ISI)
9. Siddiqui H, Healy E, Olmsted A (2017) Bot or not. In: IEEE 12th international conference for internet technology and secured transactions
10. Haidar R, Haidar R (2017) Website navigation behavior analysis for bot detection. In: IEEE international conference on data science and advanced analytics
11. Dan J, Jieqi T, Study of bot detection on Sina-Weibo based on machine learning. National key research and development program: the research on measurement method of service quality under the condition of network and information
12. Morstatter F, Wu L, Nazer TH, Carley KM (2016) Anew approach to bot detection: striking the balance between precision and recall. In: IEEE/ACM international conference on advances in social networks analysis and mining
13. Cresci S, Di Pietro R, Petrocchi M, Spognardi A, Tesconi M (2017) The paradigm-shift of social spambots: evidence, theories, and tools for the arms race. In: Proceedings of the 26th international conference on World Wide Web companion. ACM, pp 963–972
14. Cresci S, Di Pietro R, Petrocchi M, Spognardi A, Tesconi M (2017) Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling. IEEE Trans Dependable Secure Comput
15. Pedregosa (2011) Scikit-learn: machine learning in python. JMLR 12:2825–2830

16. Karatas G, Demir O, Sahingoz OK (2018) Deep learning in intrusion detection systems. In: International congress on big data, deep learning and fighting cyber terrorism. IEEE
17. Ameen AK, Kaya B (2018) Spam detection in online social networks by deep learning. IEEE
18. Lee J, Lim J, Cho W, Kim HK (2013) I know what the BOTs did yesterday: full action sequence analysis using Naïve Bayesian algorithm. IEEE
19. Chauhan SK, Goel A, Goel P, Chauhan A, Gurve MK (2017) Research on product review analysis and spam review detection. In: 4th international conference on signal processing and integrated networks (SPIN). IEEE
20. Tharwat A (2018) Classification assessment methods. Appl Comput Inf. https://doi.org/10.1016/j.aci.2018.08.003

# CYPUR-NN: Crop Yield Prediction Using Regression and Neural Networks

**Sandesh Ramesh, Anirudh Hebbar, Varun Yadav, Thulasiram Gunta, and A. Balachandra**

## 1 Introduction

For centuries, agriculture is considered to be the main and the primary culture practiced all around the globe. People in the olden days have cultivated crops in their land and hence have been accommodated to their needs [1]. Predicting the yield of the crop is a vital agricultural problem. Every single farmer constantly tries to estimate how much yield can be expected from their fields. In the past, the prediction of yield was calculated by analyzing the farmer's previous results on a particular crop. Crop yield is primarily dependent on weather conditions, pests, and the planning of harvest operation. Accurate information about the history of crop yield is a vital criterion for making decisions related to agricultural risk management. The proposed method uses regression and neural network techniques to predict the yield of paddy. These techniques have plenty of applications. Some of them are discussed below:

- Selection of crop and prediction of the yield—To aggrandize the yield of a crop, the identification and selection of the ideal crop play an important role. It is also dependent on other factors like temperature, humidity, luminescence, and external pressure that surround that crop.
- Weather forecasting—Since farmers have poor access to the Internet, they are heavily reliant on the little, yet vital information available concerning weather reports through newspapers or just pure hope. Artificial neural networks have been adopted extensively for this purpose. Newly developed algorithms have shown better results over previous conventional algorithms.
- Smart Irrigation System—The groundwater levels continue to deplete day by day, and global warming has caused drastic climatic changes. As a result, various sensor-based technologies meant for smart farming that use sensors to monitor

S. Ramesh · A. Hebbar (✉) · V. Yadav · T. Gunta · A. Balachandra
Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore 560064, India

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_17

the water level, nutrient content, weather forecast reports, and soil temperature have been introduced. Objectives of the proposed model discussed in this paper are as follows:

1. Capturing a picture of the crops to determine yield.
2. Analyzing the picture to detect diseases, if present, using neural networks.
3. Calculating values of external conditions such as pressure, humidity, and temperature.
4. Calculating accuracy levels of the probable yield
5. Indicating solution to the disease-causing pathogen, if present.

The paper is sectionalized as follows—Sect. 2 discusses a brief overview of the work carried out by previous researchers in the domain of crop yield prediction and the requirements needed for the same. Section 3 illustrates and describes the framework of CYPUR-NN, and Sect. 4 tabulates the readings obtained through experimentation of the model and methodology used. Section 5 discusses the results. Section 6 concludes the contribution of the paper with a concise summary.

## 2 Literature Survey

As per the exploration paper by Hanks [1], the creators utilized information mining procedures to take care of the issue of yield forecast. Various information mining procedures were utilized and assessed in farming for evaluating what is to come year's yield creation. Their paper additionally presents a concise investigation of harvest yield forecast utilizing multiple linear regression (MLR) method and density-based grouping strategies [2].

González Sánchez et al. [3], in their paper, thought about the prescient exactness of ML and direct relapse procedures for crop yield forecast in ten harvest datasets. Numerous straight relapse, M5-Prime relapse trees, perceptron multilayer neural systems, bolster vector relapse, and K-nearest neighbor techniques were positioned. Four precision measurements were utilized to approve the models: the root means square blunder (RMS), root-relative square mistake (RRSE), standardized mean outright mistake (MAE), and connection factor (*R*). Genuine information of a water system zone of Mexico was utilized for building the models. The outcomes indicated that M5-Prime and k-closest neighbor methods acquired the least normal RMSE mistakes (5.14 and 4.91), the most minimal RRSE blunders (79.46 and 79.78%), the most reduced normal MAE blunders (18.12 and 19.42%), and the most elevated normal connection factors (0.41 and 0.42). Since M5-Prime accomplished the biggest number of harvest yield models with the least blunders, it was an entirely reasonable instrument for monstrous harvest yield expectation in farming arranging.

In the paper proposed by Drummond et al. [4], the creators' exploration depends on understanding the connections among yield and soil properties and topographic qualities in exactness farming. A vital initial step was to recognize procedures to dependably evaluate the connections among soil and topographic qualities and harvest yield. Stepwise various direct relapse (SMLR), projection interest relapse (PPR), and a few kinds of administered feed-forward neural systems were researched trying to recognize strategies ready to relate soil properties and grain yields on a point-by-point premise inside ten individual site-years. To abstain from overfitting, assessments depended on prescient capacity utilizing a 5-crease cross-approval procedure. The neural procedures reliably outflanked both SMLR and PPR and gave insignificant forecast blunders in each site-year. A second period of the investigation included the estimation of harvest yield over different site-years by including climatological information. The ten site-long stretches of information were added with climatological factors, and expectation blunders were registered. The outcomes demonstrated that noteworthy overfitting had happened and shown that a lot bigger number of climatologically extraordinary site-years would be required in this kind of investigation.

Gandhi [5] in their paper utilized neural systems to anticipate rice creation yield and research the variables influencing the rice crop yield. The parameters considered for the investigation were precipitation, least temperature, normal temperature, greatest temperature, reference crop evapotranspiration, zone, creation, and yield for the Kharif season (June to November) for the years 1998–2002. A multilayer perceptron neural network was created. A cross-approval technique was utilized to approve the information. The outcomes indicated an exactness of 97.5% with an affectability of 96.3 and a particularity of 98.1.

Simpson [6] in their paper explored the utilization of a quick cerebellar model enunciation controller (CMAC) neural system for crop yield forecast. To begin with, expectation execution was assessed utilizing just month to month agro-met information (soil dampness, temperature, daylight). At that point, the improvement in forecast execution in the wake of fusing remote detecting information (Landsat TM) was estimated. The standard blunder was 5% when TM information was incorporated, versus 6% when TM was disregarded. The CMAC neural system applied for this investigation had recently been effectively applied in two comparable spaces: continuous cloud arrangement on Meteosat information and mineral recognizable proof with airborne obvious and infrared imaging spectrometer (AVIRIS) information.

In the paper proposed by Ji et al. [7], it was seen that by altering ANN parameters, for example, the learning rate and the number of concealed hubs influenced the precision of rice yield 3 expectations. Ideal learning rates were somewhere in the range of 71–90%. Littler informational indexes required less shrouded hubs and lower learning rates in model improvement. ANN models reliably delivered more precise yield forecasts than relapse models.

## 3 Methodology

Functional requirements are those that indicate what the CYPUR-NN system will require to ensure delivery or operation. For this project, it was vital to accumulate certain requirements that will be required to attain the targets set out. A use case analysis which was implemented resulted in the following functional and non-functional requirements.

Other features of CYPUR-NN include specific criteria to judge the operation of the CYPUR-NN system. They cover ease, security, support availability, operational speed, and implementation considerations. To be more specific, the user will find it very easy to capture photos and evaluate their crops without dire effort. The system is very secure and can be easily installed. With a response time of less than 10 s, CYPUR-NN is fast and reliable. One needs to consider the gravity of requirements captured from analyzing the use cases. This will aid in prioritizing the delivery of every single requirement. CYPUR-NN system was 4 designed by developing the NN model and mobile application simultaneously. The functioning of each is described in detail in the following subsections.

## *3.1 Convolutional Neural Network*

One of the main functions of convolutional neural networks (CNN) includes image recognition and classification. CNN image classification involves loading an image, processing it, and classifying it under categories mentioned [8]. For CNN models to train and test, each image that has been inputted will pass through serials of convolution layers with filters. Figure 1 represents the framework developed where convolutional neural network architecture and Android Application were integrated to make the system in real time.

Figure 2 explains the working mechanism of the convolutional neural network in detail. Initially, the images are collected and preprocessed to provide high-quality images for the model. In each layer, the CNN model extracts certain important features to understand and differentiate between each class.

Manually clicked photographs and stock images of the leaf images of the rice plant were used to train the model. The model was also trained to identify ailments that made their physical presence on the leaf and/or stem of the crop (Table 2). The accuracy metric was used to evaluate the algorithm's performance understandably.

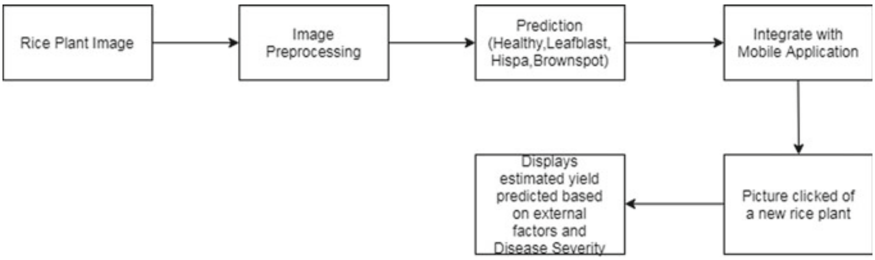



**Fig. 1** Convolutional neural network framework

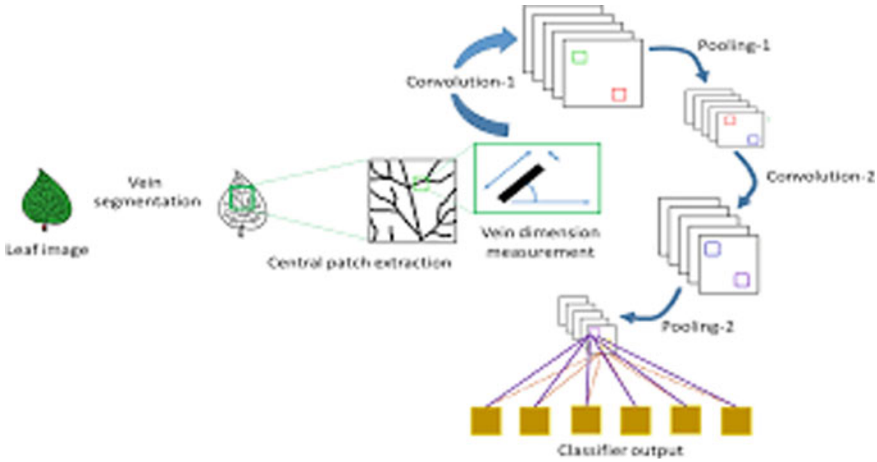



**Fig. 2** CNN framework for leaf as illustrated by González Sánchez et al. [3]

Here, the accuracy of the model was compared by the predicted label and the truth label. Categorical cross entropy was the loss function used to optimize the algorithm. It indicates as to how poorly or how conveniently a model behaves after each step of optimization. For the hidden layers, ReLU activation function (Eq. 1) was used, and for the output layer, SoftMax activation function (Eq. 2) was used. The optimizer used for the model was Adam.

ReLU activation function

$$F(x) = \max(0, x) \tag{1}$$

SoftMax activation function

$$P\left(y = j|\theta^{(i)}\right) = e^{\theta(i)} / \sum\nolimits_{j=0}^{k} e_k^{\theta(i)} \tag{2}$$

### *3.2 Neural Network-Based Mobile Application*

The neural network application takes into consideration the same parameters as that of CNN. However, to predict the yield of the crop more accurately, the application also uses sensors that measure humidity, pressure, and luminescence. To test this, paddy crops were monitored in controlled conditions. The purpose of this is being able to determine the optimum conditions for maximum crop yield. The application was also tested by replicating the conditions found in different paddy grown regions of the country. While crop yield was being measured, ailments in the crop were

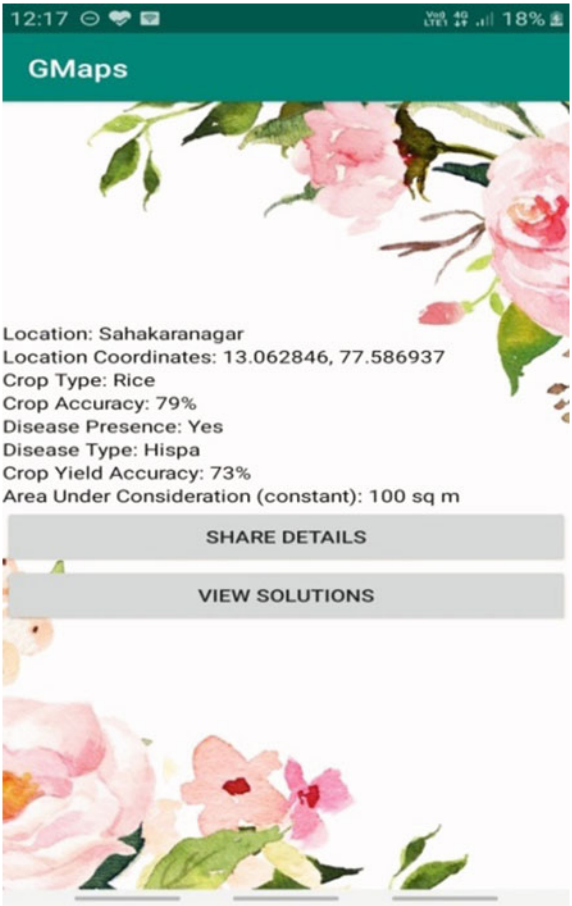



**Fig. 3** Application sensors retrieving external readings

also being monitored simultaneously. Figure 3 shows the application displaying the possibility of the disease in the rice plant.

Table 1 contains the experimental readings of simulation conditions considered optimal and suboptimal for the production of rice crops. For experimentation, the ph. levels of simulation soil for Punjab, Tamil Nadu, West Bengal, Andhra Pradesh, Bihar, and Karnataka were set to 7.8, 6, 6.5, 7, 8.4, and 5.5, respectively. These levels coincide with the data recorded by respective governments in their rice grown regions of the state. The area considered is constant for all readings, i.e., 100 $m^2$. The crops were submerged in water at all times. Temperature is considered as an important parameter. Extreme high or near low temperatures, even for short, affect crop growth. High or near low air temperature decreases the growth of shoots and also reduces the growth of the root [9]. Humidity is important to make photosynthesis possible. A good level of humidity surrounding the plant is even more vital than compared to most other crops because the plant can absorb a reduced amount of

**Table 1** Quantitative attributes of neural network application

| State conditions under simulation | Sample | Area (in sq. m) ($m^2$) | Temperature (in degree Celsius) | Humidity (%) | Pressure (in mbar) | Impact | Expected yield (in %) |
|---|---|---|---|---|---|---|---|
| Punjab | Sample P1 | 100 | 14 | 38 | 138.24 | Negative | 43 |
| | Sample P2 | 100 | 22 | 59 | 120.33 | Positive | 77 |
| | Sample P3 | 100 | 25 | 71 | 114.46 | Positive | 78 |
| | Sample P4 | 100 | 26 | 75 | 109.56 | Positive | 89 |
| | Sample P5 | 100 | 38 | 88 | 98.97 | Negative | 42 |
| Tamil Nadu | Sample TN1 | 100 | 28 | 78 | 114.67 | Positive | 89 |
| | Sample TN2 | 100 | 29 | 78 | 115.78 | Positive | 91 |
| | Sample TN3 | 100 | 33 | 80 | 99.45 | Positive | 88 |
| | Sample TN4 | 100 | 35 | 84 | 96.66 | Positive | 76 |
| | Sample TN5 | 100 | 43 | 88 | 92.34 | Positive | 71 |
| West Bengal | Sample WB1 | 100 | 29 | 78 | 112.66 | Positive | 88 |
| | Sample WB2 | 100 | 32 | 79 | 112.35 | Positive | 88 |
| | Sample WB3 | 100 | 33 | 80 | 108.67 | Positive | 91 |
| | Sample WB4 | 100 | 35 | 84 | 100.44 | Positive | 92 |
| | Sample WB5 | 100 | 40 | 81 | 99.02 | Positive | 87 |
| Andhra Pradesh | Sample AP1 | 100 | 30 | 78 | 99.65 | Positive | 88 |
| | Sample AP2 | 100 | 31 | 80 | 99.78 | Positive | 87 |
| | Sample AP3 | 100 | 35 | 82 | 91.45 | Positive | 83 |
| | Sample AP4 | 100 | 38 | 82 | 90.89 | Positive | 74 |

(continued)

**Table 1** (continued)

| State conditions under simulation | Sample | Area (in sq. m) ($m^2$) | Temperature (in degree Celsius) | Humidity (%) | Pressure (in mbar) | Impact | Expected yield (in %) |
|---|---|---|---|---|---|---|---|
| | Sample AP5 | 100 | 43 | 87 | 84.23 | Negative | 49 |
| Bihar | Sample B1 | 100 | 28 | 77 | 116.87 | Positive | 82 |
| | Sample B2 | 100 | 29 | 77 | 116.72 | Positive | 85 |
| | Sample B3 | 100 | 32 | 78 | 115.45 | Positive | 85 |
| | Sample B4 | 100 | 33 | 78 | 115.98 | Positive | 87 |
| | Sample B5 | 100 | 36 | 78 | 115.67 | Positive | 88 |
| Karnataka | Sample K1 | 100 | 23 | 61 | 109.76 | Positive | 79 |
| | Sample K2 | 100 | 24 | 61 | 109.78 | Positive | 77 |
| | Sample K3 | 100 | 28 | 66 | 101.23 | Positive | 79 |
| | Sample K4 | 100 | 33 | 75 | 99.78 | Positive | 73 |
| | Sample K5 | 100 | 35 | 77 | 90.87 | Positive | 71 |

humidity and hence has less water evaporation than compared to most plants. As plants need to transpire, the surrounding humidity saturates the leaves with water vapor. When relative humidity levels are detected to be too high or there is a lack of air circulation, a plant cannot evaporate water or extract nutrients from the soil [10]. Barometric pressure has paramount effects on water chemistry and surrounding weather conditions. It has an impact on the amount of gas that can be allowed to dissolve in water. More gas, such as oxygen, can dissolve in water under higher pressure when compared to lower air pressure [11]. The impact parameter indicates whether or not the crop germination will be successful or not. To ease the purpose of understanding, we have set a threshold level of 50% yield expectancy. Any number below this is considered suboptimal yield.

Table 2 explains the accuracy of the prediction made by the model for each class, i.e., healthy, hispa, leaf blast, and brown spot.

In the next section, alternative technique has been proposed to solve the early and reliable system to predict yield of the plant.

**Table 2** Quantitative attributes of neural network application

| Captured picture | Category | Species | Ailment condition | Accuracy |
|---|---|---|---|---|
|  | Non-leguminous plant | *Oryza sativa* | No ailment detected in given sample | 83.7858 |
|  | Non-leguminous plant | *Oryza sativa* | Ailment detected in plant sample. Possible ailment-hispa, dryness | 84.7625 |
|  | Non-leguminous plant | *Oryza sativa* | Ailment detected in plant sample. Possible ailment-fungal infection, leaf blast | 82.3554 |
|  | Non-leguminous plant | *Oryza sativa* | Ailment detected in plant sample. Possible ailment-black spots, brown spots | 85.4563 |

## *3.3 Multiple Linear Regression*

Multiple linear regression is a statistical technique, where multiple independent variables are used to predict one output variable. In this proposal, our 6 independent features consist of the area of the land used by the farmer to cultivate the rice plant, the state in which he lives in and the season based on which he/she wants to know the yield of the plant. The output feature consists of the estimated yield based on the inputs given by the user. The formula of the multiple linear regression is as follows (Fig. 4):

$$y_i = b_0 + b_1 * xi_1 + b_2 * xi_2 + b_3 * xi_3 + \cdots + b_n * xi_n$$

where

$i$ $\quad$ $n$ observations

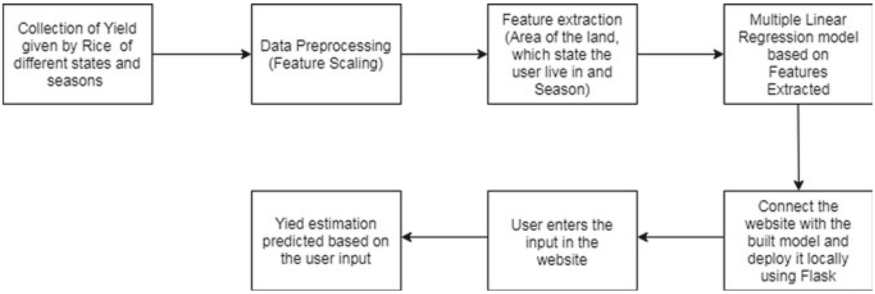



**Fig. 4** Framework for multiple linear regression model

$y_i$ dependent or target variable
$x_i$ independent variables
$b_0$ $y$—intercept
$b_n$ slope coefficients.

In this model, eleven independent features were considered since features such as state and seasons were one-hot-encoded along with the area of the land before fitting them to the model. The state feature mainly consisted of Andhra Pradesh, Karnataka, Kerala, Pondicherry, and Tamil Nadu. These features were considered keeping data distribution and curse of dimensionality in mind. The season feature mainly consisted of Autumn, Kharif, Summer, Rabi, and Winter. To package the model as a product, a simple Web interface was created which will be discussed in the upcoming section.

In the multiple linear regression model, we have established a relationship between the crop yield and the season. Although the size of the land does become an important factor as they are positively related to each other, we cannot expect the same yield throughout the year. Thus, keeping this in mind we have considered Kharif season as the best rice-growing season as the yield predicted in this season is higher compared to other seasons when other features such as area and state were kept constant.

### *3.4 Web Application for Multiple Linear Regression*

Any machine learning model is incomplete if it is not packaged with a user interface and makes it available to the consumers. Keeping that in mind a simple Web interface was designed and hosted locally using Python micro-service known as Flask. Figure 5 is the Web interface used for the regression model.

Initially, a user enters the total area in which he/she enters the total area (in hectares) in which the person wishes to grow rice plants. Next, the user is going to select the state he/she lives in and finally the season in which he wishes to choose to estimate the yield it might give. After the predict button is clicked, the model is

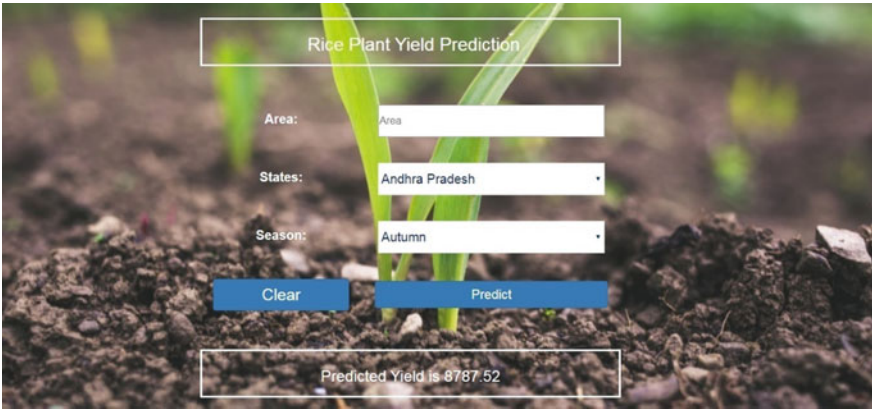



**Fig. 5** Web application for multiple linear regression

going to consider the input values, processes it, and finally, it predicts the estimated yield which is then displayed to the user.

# 4 Results and Discussions

## *4.1 Results for Convolutional Neural Network*

For the trained convolutional neural network, the training set accuracy and test set accuracy were 86.37% and 83.87%, respectively. Figures 6 and 7 show the model performance in terms of accuracy and loss for 180 epochs.

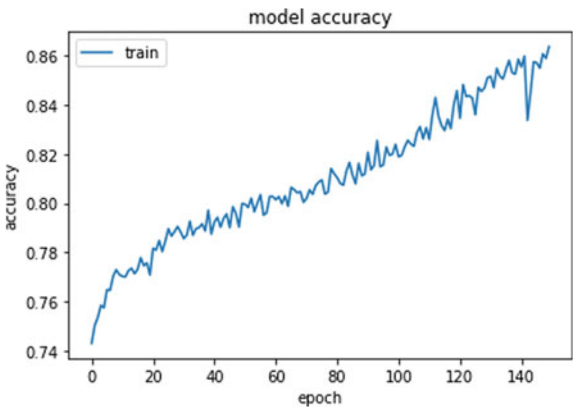



**Fig. 6** Epoch v/s accuracy graph for CNN model

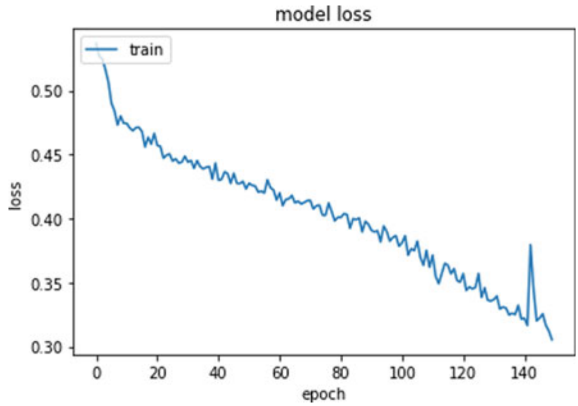



**Fig. 7** Epoch v/s loss graph for CNN model

From Figs. 6 and 7, we can see that as the iterations in the model increase, the accuracy of the model increases while the loss decreases accordingly, thus the model is improving its performance at each iteration. The attributes obtained by experimenting with the neural network application are tabulated in Table 1. Punjab experiences a minimum of 14 °C and a maximum of 38 °C. It was observed that at these two temperatures, the impact was negative and the expected yield was suboptimal. Sample P4 with an ideal temperature of 26 °C, the humidity of 75%, and an atmospheric pressure of 109.56 was expected to produce the highest yield of 89%. Tamil Nadu experiences a minimum of 28 °C and a maximum of 43 °C. While these conditions are considered very warm, the growth of rice remained prominent throughout changes in conditions. Sample TN2 resulted in the highest 91% yield expectancy. The corresponding conditions were 29 °C, 78% humidity, and 115.78 mbar. The lowest yield expectancy was for Sample TN5, with 71%. West Bengal experiences a minimum of 29 °C and a maximum of 40 °C. Like Tamil Nadu, West Bengal too experiences very warm temperatures. Despite this, West Bengal has the highest rice yield compared to all the other states in the table below. With 92% yield expectancy, the conditions of WB4 are considered ideal for rice growth. Andhra Pradesh, another major rice-growing state in the country, shows satisfactory yield expectancy. However, when the temperature soars to 43 °C and pressure drops to 84.23 mbar, the yield too drops to a mere 49%. Bihar too had consistent results with all samples averaging between 82 and 88%. Karnataka on the other hand experiences almost moderate climatic conditions throughout the year, with the highest being 35 °C and the lowest being 23 °C. The optimum conditions in Karnataka are considered to be similar to Sample K3 (28 °C, 66% humidity, and 101.23 mbar atmospheric pressure).

## 4.2 *Results for Multiple Linear Regression*

The proposal via multiple linear regression is equally promising.

The root mean squared value obtained for the model is 0.343. The intercept value is around 0.874.

The coefficient values for each attribute are as follows:

| Features selected for the model | Coefficients of the features selected |
|---|---|
| Area (in hectares) | 1.019 |
| Andhra Pradesh | 0.035 |
| Karnataka | −0.082 |
| Kerala | −0.309 |
| Pondicherry | 0.095 |
| Tamil Nadu | 0.260 |
| Autumn | 0.041 |
| Kharif | −0.157 |
| Rabi | −0.029 |
| Summer | 0.111 |
| Winter | 0.034 |

Based on the coefficients obtained following will be our multiple linear regression formula which can be used to estimate yield in the future.

$$\begin{aligned}\text{Estimated Yield} &= 0.874 + 1.019 * \text{Area}\\ &\quad + 0.035 * \text{Andhra Pradesh} + (-0.082) * \text{Karnataka}\\ &\quad + (-3.09) * \text{Kerala} + 0.095 * \text{Pondicherry}\\ &\quad + 0.260 * \text{Tamil Nadu} + 0.041 * \text{Autumn}\\ &\quad + (-0.157) * \text{Kharif} + (-0.029) * \text{Rabi}\\ &\quad + 0.111 * \text{Summer} + 0.034 * \text{Winter}.\end{aligned}$$

## 5 Conclusion

This paper has successfully demonstrated the use of convolutional neural networks and multiple linear regression techniques in the development of CYPUR-NN model on a dataset consisting of various parameters related to the obtaining of expected rice yield. The model was also experimented in a simulated environment to replicate climatic conditions in various rice-growing states. The results were very realistic in nature. The agricultural sector is considered to have an important role; also, it is very important in the global economy country in the world. The use cases of machine learning have gone on to become trending, and large-scale advancements in technology have been widely utilized in modern-day agricultural technology. Artificial intelligent techniques and methodologies are currently being used extensively in the agricultural sector as a single purpose to aggregate the accuracy and to identify solutions to the problems. A practical usage of artificial intelligent (AI) based on convolutional neural networks (CNN) application in a plethora of fields shows that CNN-based machine learning scheme is open to change and can be implemented on an agricultural field.

## References

1. Hanks RJ (1974) Model for predicting plant yield as influenced by water use 1. Agron J 66(5):660–665
2. Ramesh D, Vardhan BV (2015) Analysis of crop yield prediction using data mining techniques. Int J Res Eng Technol 4(1):47–473
3. González Sánchez A, Frausto Solís J, Ojeda Bustamante W (2014) Predictive ability of machine learning methods for massive crop yield prediction
4. Drummond ST, Sudduth KA, Joshi A, Birrell SJ, Kitchen NR (2003) Statistical and neural methods for site–specific yield prediction. Trans ASAE 46(1):5
5. Gandhi N, Armstrong LJ (2016) A review of the application of data mining techniques for decision making in agriculture. In: 2016 2nd international conference on contemporary computing and informatics (IC3I), IEEE, pp 1–6
6. Simpson G (1994) Crop yield prediction using a CMAC neural network. In: Image and signal processing for remote sensing, vol 2315. International Society for Optics and Photonics, pp 160–171
7. Ji B, Sun Y, Yang S, Wan J (2007) Artificial neural networks for rice yield prediction in mountainous regions. J Agric Sci 145(3):249–261
8. Guo WW, Xue H (2014) Crop yield forecasting using artificial neural networks: a comparison between spatial and temporal models. Math Problems Eng 2014
9. Panda SS, Ames DP, Panigrahi S (2010) Application of vegetation indices for agricultural crop yield prediction using neural network techniques. Remote Sens 2(3):673–696
10. Chen C, McNairn H (2006) A neural network integrated approach for rice crop monitoring. Int J Remote Sens 27(7):1367–1393
11. Jiang D, Yang X, Clinton N, Wang N (2004) An artificial neural network model for estimating crop yields using remotely sensed information. Int J Remote Sens 25(9):1723–1732

# Static and Dynamic Human Activity Detection Using Multi CNN-ELM Approach

**Shilpa Ankalaki and M. N. Thippeswamy**

## 1 Introduction

Human Activity Recognition has its foremost impact on applications like smart health care and assisting elders, and it is necessary for humankind to assist human's daily life by recording the human movements from the sensors that are incorporated into the human body in the form of smart devices. Selecting proper methods for analyzing sensor data to make the correct decision is one of the challenging tasks. In this view, deep learning has immense benefits in HAR. HAR can be accomplished in two ways are Vision-Based Approach and Sensor-Based Approach. The former approach makes use of visual cameras to capture the image/video and data will be processed using image processing techniques and; analyzed by using machine learning and deep learning techniques whereas sensor-based HAR can be accomplished by using non-visual sensors. Vision-based HAR has its limitation in applications due to its consideration of secrecy concerns of mounting vision sensors in private space [1]. Non-visual sensors based HAR has the advantage of generality because sensors can be easily embedded in almost all devices including the human body. The primary difficulty of Sensor-based HAR is the representation of information captured by various sensors. Traditional classifiers exhibit limited performance for HAR because of the process of extraction of handcrafted features. This drawback is overcome by using deep learning techniques that provide the facility of automatic feature extraction from given data.

S. Ankalaki (✉) · M. N. Thippeswamy
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore 560064, India

M. N. Thippeswamy
e-mail: thippeswamy.mn@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_18

There are a couple of approaches under sensor-based HAR, those are knowledge-driven approach and data-driven approach [2]. A knowledge-driven approach constructs activity models by taking advantage of complete prior knowledge in the area of interest. The data-driven approach employs publically available datasets to study the activity models by applying machine learning and deep learning techniques. The proposed work emphasizes data-driven solutions to HAR, also discussed the existing barriers of their application on UCI-HAR dataset. This work focuses on the recognition of human static and dynamic activities of the UCI-HAR dataset.

Followings are the objectives and key contributions of this proposed work:

- Proposed the hierarchical 1D-CNN approach for Classification of Human Activities into Static and Dynamic activities
- Proposed the novel hybrid 1D-CNN-ELM approach for classifying the static activities into sitting, standing, and lying; and dynamic activities into walking, walking_upstairs, and walking_downstairs.
- Perfomance evaluation of 1D CNN-ELM approach for classification of static and dynamic activities using precision, recall, confusion matrix, and accuracy measures.

The rest of the paper is organized as follows: Sect. 2 briefs about understanding the human activity and human activity recognition; a literature survey of deep learning approaches for sensor-based HAR. Section 3 provides the details of the proposed methodology employed to accomplish the above-stated objectives. Section 4 provides the experimentation results and analysis of the proposed work.

## 2 Related Work

This section provides the relevant literature about the understanding of human activity, human activity recognition, and deep learning methodologies for sensor-based HAR.

### *2.1 Understanding Human Activity*

Human activities are the order of human movements operated by an individual over a period of time in a given ambient. In the view of senor based HAR, the activity can be well-defined as the set of actions where each action consists of a sequence of events. Events are interpreted as a sequence of data generated by various sensors records, whereas usually sensors are incorporated in human bodies but in advanced HAR sensors are incorporated in the environment as well [3, 4]. The mathematical representation of activity definition is represented by the Eqs. (1)–(2).

$$A = (A_i)_{i=1}^{m} \tag{1}$$

In Eq. (1), A represents the existing activity set which is inclusive of ‘m’ number of various activities. The sequence of data that is captured by the sensor for a given period of time is represented by an Eq. (2).

$$S = \{r_1, r_2, \ldots r_t, \ldots, r_n\} \tag{2}$$

where $r_t$ represents the reading of the sensor at time $t$.

The objective here is to construct the model that predicts the series of activities that belong to set $A$ depending on the senor reading $S$.

## 2.2 Human Activity Recognition

In our everyday life, people perform 2 kinds of traditional physical activities those are namely, static and dynamic human activities. Sitting, standing and sleeping are some of the activities which are inclusive of static activities. Human physical activities can also be categorized as atomic, simple, and complex activities [5]. Atomic activities are static activities, Standing is one such example. Simple activities are a systematic sequence of static activities performed within a specified time interval, Walking is one such activity that belongs to the simple activity category. Complex activities are an assortment of more than one simple activity that takes place at a specified time. Dancing is an example of a complex activity. Researchers like Lara and Labrador [6] summarized seven types of activities recognized by the HAR system in their literature. Those activity groups are namely: Daily activities, Transportation, Fitness, Phone usage, Military, Ambulation, and Upper Body activities.

Sensor-based approaches for human activity recognition can be accomplished by incorporating the sensors in the human body and environment. An approach of incorporating the sensors in the human body provides sensor data by which the system can identify individual activities like walking, dancing, skipping, cooking, etc. without considering the location. Whereas sensors incorporated in an environment gathers information about the locations where activities are taken place [7].

Sensor-based approaches are created a successful path for recognition of human activities ranging from static to dynamic and simple to complex human activities. One such work is accomplished by AROMA which recognizes both simple and complex activities together [8]. Researchers [2] successfully created a smart environment in UJAmI Smart Lab by incorporating binary sensors in the kitchen area, bedroom area, Door, Laundry basket, sofa. The sensor-based HAR can be accomplished by using machine learning and deep learning approaches. Deep learning approaches for sensor-based HAR are discussed in the subsequent section.

### *2.3 Deep Learning Methodologies for Sensor-Based HAR*

The recent literature provides evidence that the deep learning methods outperform machine learning algorithms for sensor-based HAR because of its unique characteristics of automatic feature extraction. One such survey has been conducted in [9] which provides an extensive study about deep learning approaches for sensor-based HAR and also focused on the numerous ways to address the challenges of HAR. The extensive study has been carried out by the researchers of [10] about feature learning using a convolutional neural network for HAR. The experiment was conducted on the datasets like DCASE 2017 dataset, Extrasensory dataset, UCI-HAR dataset, and real-world Extrasensory dataset; and analyzed the performance of various architectures of CNN and fine-tuning the CNN by changing the values of its hyperparameters. Apart from the recognition of Daily human activities, researchers also contributed to the recognition of athletic tasks using deep neural networks [11]. This task has been accomplished by using the hybrid approach by combining CNN and LSTM. The experiment has been carried out from the sensor data collected from 417 athletes where every athlete has 13 athletic movements. The work discussed in [12] proposed a novel approach for HAR which adopts pose reconstruction dataset AMASS along with virtual IMU data. This dataset is trained by using the collective framework of CNN along with an unsupervised penalty for HAR. The experiment was conducted by using the hybrid architecture of CNN-RNN on Opportunity dataset, PAMAP2 dataset, and Daphnet Gait dataset and concluded that bidirectional LSTM outstrips the other algorithms on the opportunity dataset [13].

One of the missing features in formerly described contributions about the deep learning architectures is the grouping of activities present in the dataset into static and dynamic activities. The deep learning approaches described in this section classifies the activities into the individual classes without considering whether it is a static and dynamic activity. The proposed work focuses on the hierarchical and hybrid 1D CNN-ELM approach for activity recognition and performance evaluation of the classification of each activity using various metrics.

## 3 Data and Methodology

The proposed work aims to assess the performance of the hybrid CNN-ELM approach and includes performance analysis by fine-tuning the hyperparameters. Figure 1 depicts the detailed pictorial representation of the proposed methodology.

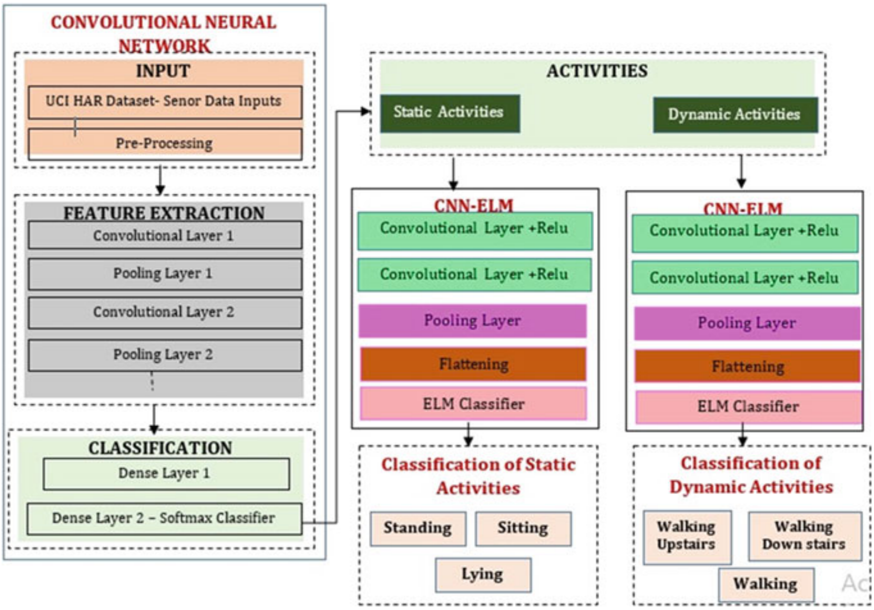



**Fig. 1** Hierarchical multi CNN-ELM approach for classification of static and dynamic activities

## 3.1 Input Dataset

The proposed methodology has employed the UCI-HAR dataset to accomplish sensor-based HAR. The aforementioned dataset was constructed by [14], the dataset built by the activities performed by 30 individuals. Every individual has accomplished the following activities using wearable sensors, those are, lying, standing, sitting, walking_upstairs, walking, and walking_downstairs.

## 3.2 Pre-processing

In this stage, Data cleaning has been carried out on the UCI-HAR dataset by examining the dataset for missing values, null values, and balanced data distribution for each target class.

## 3.3 Feature Learning Using CNN

The proposed work employed, CNN for instinctive feature extraction from raw data of sensors which eliminates the step of handcrafted feature extraction. Use of CNN

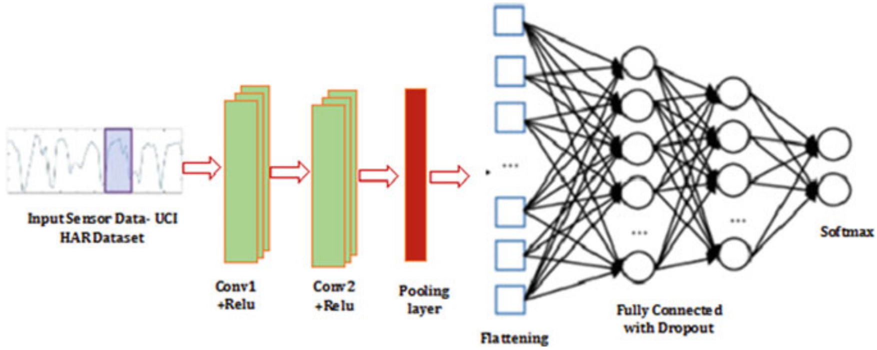



**Fig. 2** Structure of 1-D CNN employed in proposed work

simplifies the activity recognition by removing the process of Handcrafted Feature extraction, which usually requires domain-specific experts to recognize preferable and distinct features [10]. The structure of CNN includes a set of layers namely, Convolutional pooling and dense layer. CNN accomplishes the task of feature extraction using a convolutional layer with varying the number and size of filters in each layer and non-liner feature extraction. Downsampling is accomplished by using the pooling layer and classification is accomplished by using a dense layer with a softmax activation function.

Usually, CNN's are categorized based on the dimension of kernels used in the convolutional layer those are, 1-D, 2-D, and 3-D CNN. The proposed work employed 1-D CNN as it uses the sensor data for feature extraction. Figure 2 describes the structure of CNN used in this work.

Figure 2 describes the structure of CNN employed in this work. Input is 1-dimensional sensor data, this raw data is passed through the two sets of the convolutional layer. In this stage, the features are extracted at the convolutional layer by performing the convolutional operations on raw sensor data using 1-dimensional kernels (filters). The nonlinear activation functions like Rectified Linear Unit (Relu) or Leaky Relu functions are applied to the resultant features obtained by the convolutional layer to make features highly nonlinear. The process of feature extraction using a one-dimensional kernel is depicted in Fig. 3.

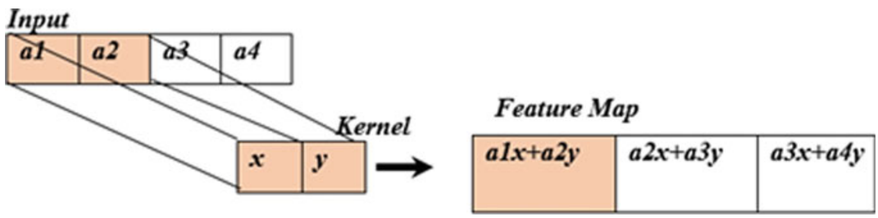



**Fig. 3** Feature extraction using 1-D kernel in convolutional layer. *Source* [15]

Downsampling has been accomplished by using max or average pooling. The resultant features are flattened into a vector and it is considered as trainable learning which will be given as input to the fully connected network. The commonly used approach for the fully connected network is Multi-Layer Perceptron. The proposed work employed Hierarchical Multilevel CNN, in which Multi-Layer Perceptron (MLP) has been used at the root level to classify the activities into static and dynamic activities followed by two CNN –Extreme Learning Machine to classify sub-activities of static and dynamic activities.

### *3.4 Extreme Learning Machine (ELM)*

The iterating nature which causes more computational effort is one of the foremost shortcomings of MLP [16]. In order to overcome this issue, the proposed work employed ELM instead of MLP at a fully connected neural network. The ELM was originally proposed by Huang et al. [17]. The procedure of ELM [16–18] is given in Table 1.

## 4 Results and Analysis

In this work, the hierarchical multi CNN-ELM approach has been used for classifying the activities of the UCI-HAR dataset. The first level involves the classification of

**Table 1** Algorithm for extreme learning machine

**Input:** Training Dataset ($A_i$,$Y_i$ ) ($A_i$ stands for attributes, and $Y_i$ is target class label),
**Output:** Prediction of class labels
**Steps:**
Step 1: Initialization of arbitrarily engendered weights wi and biases bi from input to hidden layer
Step 2: Compute the output matrix H for hidden layer, where H = {$h_{ij}$} (i = 1,…, N and j = 1, …, M) and

$$h_{ij} = g(w_j * a_j + b_j) \tag{3}$$

Step 3: Compute linear transfer function of output layer as showed in Eq. (4)

$$\sum_{i=1}^{n} \beta_i g(w_j * a_j + b_j) = T_j \tag{4}$$

Where, n – Number of hidden neurons
a - Input Vector
w and T– Weight and Output Vector respectively.
$\beta$ and $g(y)$ – Transfer function of Output and Hidden layer
The matrix representation of Eq. (4) is showed in Eq. (5)

$$H\beta = T \tag{5}$$

Where,

$$H = [g(w_1 * a_1 + b_1) \ldots g(w_n * a_1 + b_n) \ldots g(w_1 * a_N + b_1) \ldots.. g(w_n * a_N + b_n)]$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \qquad T = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$$

*Step 4:* Eq. (6) represents the computation of the weights matrix for output layer

$$\beta = H \dagger T \tag{6}$$

where H† denotes the MP generalized inverse of matrix H.

human activities into static and dynamic activities using CNN with MLP classifier whereas the second level has the combination of 2 CNN-ELM classifiers to classify the static and dynamic activities into its respective sub-activities.

### *4.1 Classification of Static and Dynamic Human Activities*

In the UCI-HAR dataset, the magnitudes are represented by features like tBodyAccMag, tGravityAccMag, tBodyAccJerkMag, tBodyGyroMag, and tBodyGyroJerkMag.

Figure 4 represents the graph of static and dynamic activities with respect to the mean of body acceleration feature, i.e., tBodyAccMag_mean. From the analysis, we can conclude that static and dynamic activities are completely different and classification of these activities is accomplished by structural design of CNN depicted in Fig. 5. This structural design of CNN represents the first level binary classifier which classifies the activities into static and dynamic activities. The architecture consists of 2 successive convolutional layers with 32 filters in each layer; and size of the kernel 7 and 3 respectively followed by Max pooling (pool_size = 3) and flattening. The Relu activation function has been used in all convolution layers. The classification of activities has been accomplished by dense layer with a softmax activation function. The training has been performed for 20 epochs and with batch size 32. Categorical Cross entropy is employed for loss calculation and Adam optimizer with a learning rate of 0.004 is used for optimization. Figures 6 and 7 depicts the training and validation accuracy and training and validation loss of first level CNN architecture respectively.

In the second level, classification of static activities into sit, stand, and lying down; and moving activities into walking, walking upstairs, and walking downstairs is accomplished by using two 3-class CNN-ELM Classifier which is depicted in Fig. 8. The architecture consists of 2 successive convolutional layers with a number of filters 64 and 32; and the size of the kernel 7 and 3 respectively followed by Max

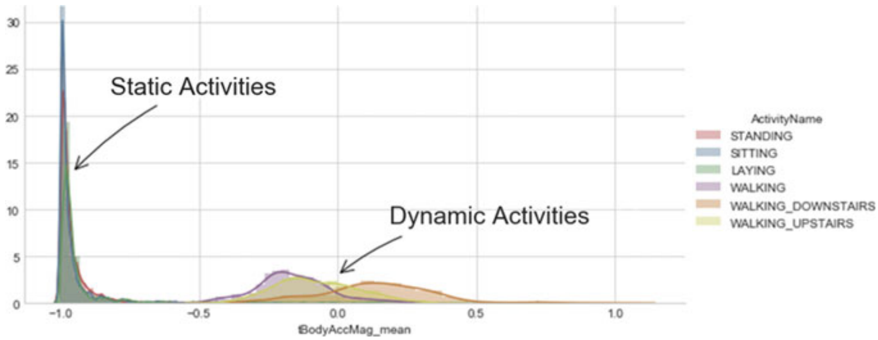



**Fig. 4** Analysis of static and dynamic activities with respect to tBodyAccMag_mean feature

| Layer (type) | Output Shape | Param # |
|---|---|---|
| conv1d_1 (Conv1D) | (None, 126, 32) | 896 |
| conv1d_2 (Conv1D) | (None, 124, 32) | 3104 |
| dropout_1 (Dropout) | (None, 124, 32) | 0 |
| max_pooling1d_1 (MaxPooling1 | (None, 62, 32) | 0 |
| flatten_1 (Flatten) | (None, 1984) | 0 |
| dense_1 (Dense) | (None, 50) | 99250 |
| dense_2 (Dense) | (None, 2) | 102 |

**Fig. 5** CNN architecture to classify activities into static and dynamic activities

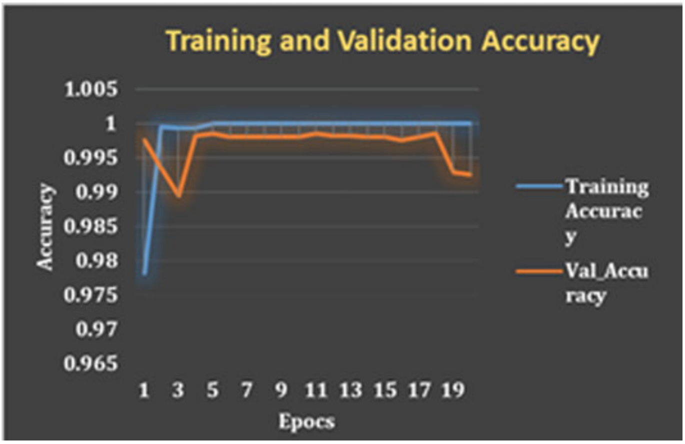



**Fig. 6** Training and validation accuracy

pooling (pool_size = 3) and flattening. The Relu activation function has been used in all convolution layers. The classification of activities has been accomplished by ELM classifier with 20 hidden neurons. The training has been performed for 20 epochs and with batch size 32. Categorical Cross entropy is employed for loss calculation and Adam optimizer with a learning rate of 0.004 is used for optimization.

The confusion matrix has been calculated for both classification static and dynamic activities and the same is showed Tables 2 and 3.

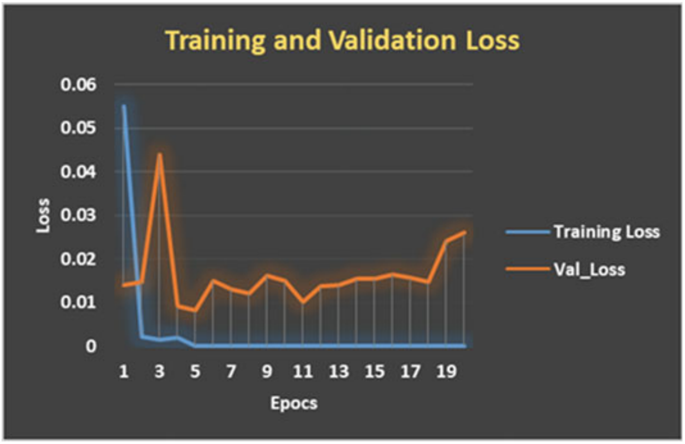



**Fig. 7** Training and validation loss

| Layer (type) | Output Shape |
|---|---|
| conv1d_1 (Conv1D) | (None, 122, 64) |
| conv1d_2 (Conv1D) | (None, 120, 32) |
| dropout (Dropout) | (None, 120, 32) |
| max_pooling1d | (None, 40, 32) |
| flatten (Flatten) | (None, 1280) |
| ELM Classifier | (10, Tanh) |

**Fig. 8** Architecture of CNN-ELM classifier

**Table 2** Confusion matrix for static activities

| Actual class | Predicted class | | | | Recall (%) |
|---|---|---|---|---|---|
| | | SIT | STAND | LYING | |
| | SIT | 534 | 3 | 0 | 99.44 |
| | STAND | 0 | 450 | 41 | 91.64 |
| | LYING | 0 | 27 | 505 | 94.04 |
| Precision (%) | | 100 | 93.75 | 92.49 | 95.44 |

**Table 3** Confusion matrix for dynamic activities

| Actual class | Predicted class | | | | Recall (%) |
|---|---|---|---|---|---|
| | | Walk | Walking_upstairs | Walking_downstairs | |
| | Walk | 486 | 0 | 10 | 97.98 |
| | Walking_upstairs | 3 | 415 | 3 | 98.57 |
| | Walking_downstairs | 3 | 2 | 465 | 98.93 |
| Precision (%) | | 98.78 | 99.52 | 97.28 | 98.48 |

## 5 Conclusion

The proposed work focuses on the hierarchical multilevel CNN-ELM classifiers for the recognition of static and dynamic activities. This work concludes that the differentiation of static activities, i.e., standing and lying was a difficult task as it is evinced by the comparatively large number of the misclassified instances (41 and 27 respectively). The precision and recall values are better for dynamic activity recognition compares to static activities, which concludes that the classification of dynamic activities is more accurate than the static activities. This classification achieves accuracy of 95.44% for static activities and 98.48% for dynamic activities. The proposed method accomplished the overall accuracy of 96.86% on the UCI-HAR dataset. This work can be extended by fine-tuning the parameters of CNN and by varying the ELM parameters like activation functions and the number of hidden neurons.

## References


1. Bevilacqua A et al (2019) Human activity recognition with convolutional neural networks. In: Lecture notes in computer science, pp 541–552
2. Irvine N, Nugent C, Zhang S, Wang H, Ng WWY (2020) Neural network ensembles for sensor-based human activity recognition within smart environments. Sensors 20:216
3. Wang J et al (2019) Deep learning for sensor-based activity recognition: a survey. Pattern Recogn Lett 119:3–11
4. Bulling A, Blanke U, Schiele B (2013) A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys
5. Nehal S, Abu-Elkheir M, Atwan A, Hassan S (2018) Current trends in complex human activity recognition. J Theor Appl Inf Technol 96(14):4564–4583
6. Lara OD, Labrador MA (2013) A Survey on human activity recognition using wearable sensors. IEEE Commun Surv Tutor 15(3):1192–1209
7. Yang J (2009) Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In: Proceedings of the 1st international workshop on interactive multimedia for consumer electronics, Beijing, China, 23 October 2009, pp 1–9
8. Peng L et al (2018) AROMA: a deep multi-task learning based simple and complex human activity recognition method using wearable sensors. IMWUT 2: 74:1–74:16
9. Chen K, Zhang D, Yao L, Guo B, Yu Z, Liu Y (2020) Deep learning for sensor-based human activity recognition: overview, challenges and opportunities. arXiv:2001.07416v1 [cs.HC] 21 Jan 2020

10. Cruciani F, Vafeiadis A, Nugent C et al (2020) Feature learning for human activity recognition using convolutional neural networks. CCF Trans Pervasive Comp Interact 18–32
11. Clouthier A, Ross G, Graham R (2020) Sensor data required for automatic recognition of athletic tasks using deep neural networks. Front Bioeng Biotechnol 7:473. https://doi.org/10.3389/fbioe.2019.00473
12. Xiao F, Pei L, Chu L, Danping Z, Yu W, Zhu Y, Tao L (2020) A deep learning method for complex human activity recognition using virtual wearable sensors. arXiv:2003.01874v2 [cs.CV] 6 Mar 2020
13. Hammerla NY, Halloran S, Plötz T (2016) Deep, convolutional, and recurrent models for human activity recognition using wearables. IJCAI. AAAI Press, pp 1533–1540
14. Anguita D et al (2013) A public domain dataset for human activity recognition using smartphones. ESANN
15. Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
16. Jayaweera CD, Aziz N (2018) Development and comparison of extreme learning machine and multi-layer perceptron neural network models for predicting optimum coagulant dosage for water treatment. IOP Conf Ser J Phys Conf Ser 1123
17. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine theory and applications. Neuro Comput 70:489–501
18. Silva DNG, Pacifico LDS, Ludermir TB (2011) An evolutionary extreme learning machine based on group search optimization. In: 2011 IEEE congress of evolutionary computation (CEC), New Orleans, LA, pp 574–580

# Health Assistant Bot

**Nikhil Kishore Nayak, G. Pooja, Ramya Ravi Kumar, M. Spandana, and P. Shobha**

## 1 Introduction

Chatbots are computer applications based on artificial intelligence (AI), which conduct conversations via textual methods to respond to the queries of the users. Chatbots utilize machine learning and natural language processing (NLP) to understand the user input and provide the required response. The proposed system is developed by Rasa which is an open-source machine learning framework for building chatbots. We have named this bot as Zara. The proposed system is also trained to answer questions and dispel doubts that people may have about the recent coronavirus pandemic COVID-19 which is causing chaos all around the world. It easy to use, provides a good customer satisfaction, and saves time. It makes sure that the necessities of the user are taken care of such as booking online appointments and ordering medicine online. Additionally, the bot empowers the user by providing a comprehensive list of exercises as per their choice, monitoring their medication schedule through reminders, finding the nearby healthcare centers on the map, enabling SOS during emergency. Its features like healthy recipes suggestions and solutions to medical problems motivate the users to follow their routine in a better way. Giving users the choice of language exhilarates the bot to be more user-friendly and beneficial to a larger audience.

N. K. Nayak · G. Pooja · R. R. Kumar · M. Spandana · P. Shobha (✉)
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India
e-mail: shobha.p@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_19

## 2 Literature Survey

The paper [1] proposed an idea to develop a chatbot that can predict diseases using apriori algorithm by taking symptoms as input provided by the user. This is achieved by incorporating various machine learning algorithms into the proposed system. They have also implemented recurrent neural network algorithms thereby making it simple to use. The user inputs are processed by natural language processing (NLP), the input being either speech or text. Speech-to-text conversion and vice versa are handled by Google Speech Recognition API. The Google APIs can also translate several other languages into text. In paper [2], a system was presented in which the users ask questions to the bot related to therapeutic doses by voice. The system gets the output from the API and displays the name of the respective drug. The bot communicates with the user using NLP. The analytics are performed by the support vector machine (SVM) algorithm, and the system predicts the disease based on the symptoms. The user can get the corresponding answer displayed on the Android application and refers to the medical treatment for the disease provided by the bot. The paper [3] describes the advancement of a chatbot for medical students, which uses the open-source AIML-based chatter bean. This AIML-based chatbot is streamlined to change over regular language inquiries into applicable SQL queries. A total of 97 question samples was collected. The questions were divided into different categories based on their types. The resulting categories were ranked according to the number of questions in each category. Paper [4] proposes a framework that is a Web application with a chatbot embedded in it. The system includes symptoms checker and also notifies the patients about the nearby doctors available. The bot also provides health tips. The system makes use of artificial intelligence and natural language processing (NLP) as the technology stack. The system is based on retrieval-based model on NLP where the bot is trained on a set of possible questions and answers that can be asked by the user. Paper [5] proposed a system for patients to get instant help in cases of an emergency situation using medical chatbot and SOS functionality. The SOS operation usage can be automated or manually be performed which can be used to send an instant message to the emergency contact number. The message contains details of the patient location. The system also enables the Google Map feature to view the nearby doctors in that particular location through Google Map API. Also, the doctors can view the patient details and his medical history in the application. In paper [6], a system was proposed that described a text-based healthcare chatbot (THCB) system that was designed to support patients and doctors. An analysis of the intervention group from an ongoing RCT indicated that the implemented THCB, which played the role of a peer character, engaged patients to a notable extent for more than four months. Furthermore, more than 99.5% of the interactive turns were driven by THCB which underlines the scalability of THCB.

## 3 Proposed System

The chatbot Zara is a virtual healthcare companion. Zara is deployed on Telegram. The bot is equipped with machine learning-based dialogue management and interactive learning. This has been achieved by using the Rasa open-source machine learning framework. The bot also identifies the tasks which are not designed for its purpose and gives the appropriate responses. It has several applications as listed below:

A. *Language preferences*
The bot supports all its functionalities in three different languages—English, Hindi, and Kannada. The bot can accept queries in any of these languages and respond appropriately which is more convenient for the users who prefer conversing in their native languages.

B. *Home remedies*
Zara provides a set of home remedies for non-life-threatening diseases like common cold, flu, headache, etc. (Fig. 1).

C. *SOS*
Zara is also provided with SOS facility. The SOS operation sends a text message to the emergency contact during the hour of the need. This message contains user location with text quoting "Emergency."

D. *Exercise recommendation*
The exercise recommendation provides different options to user such as daily fitness, cardio, yoga, and also health-specific. The bot acts as a virtual personal trainer for the user. Zara can also locate nearby gyms and yoga instructors. Figure 2 shows the exercise recommendations provided by the bot.

E. *Scheduling of reminders*
Zara keeps a record of the patient's medication schedule and reminds the same through notifications. The user has to schedule the reminders. Zara sends a notification alert quoting the name of medicine at the particular user-defined time.

F. *Locating hospitals nearby*
The user can find the nearby healthcare centers and pharmacies with the help of bot. The bot directs the user to the Google Map application.

G. *Nutrition food recipes*
The user can also use the service of the bot for obtaining nutritional food recipes. The users have to give certain preferences like ingredient, calorie range, and their diet type.

H. *Book online appointment*
The user can also book health appointments online through the bot by selecting their preferences such as specialization and appointment time as shown in Fig. 3. The user is redirected to the specific Web site for the same.

I. *To-Do List*
Zara can maintain a list of the users' daily events and notifies about the same on user request.

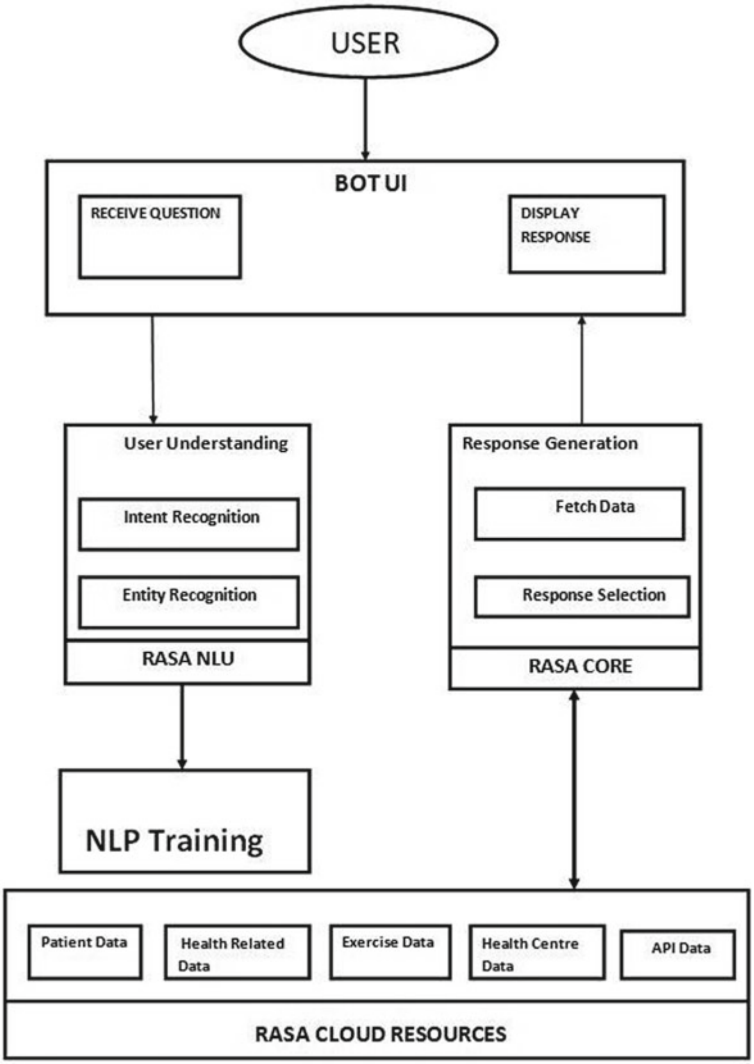



**Fig. 1** Architectural design of the proposed system

J. *Order medicine online*
In cases where the user has to place orders of medicines online, the bot redirects them to the online medicine purchase Web site.
K. *Queries regarding COVID-19 and news updates*

Zara can provide answers to all kind of queries regarding COVID-19. It informs users regarding the risk, best practices, and advisories related to spread of COVID-19 all over the world. It also gives the current corona cases count throughout the

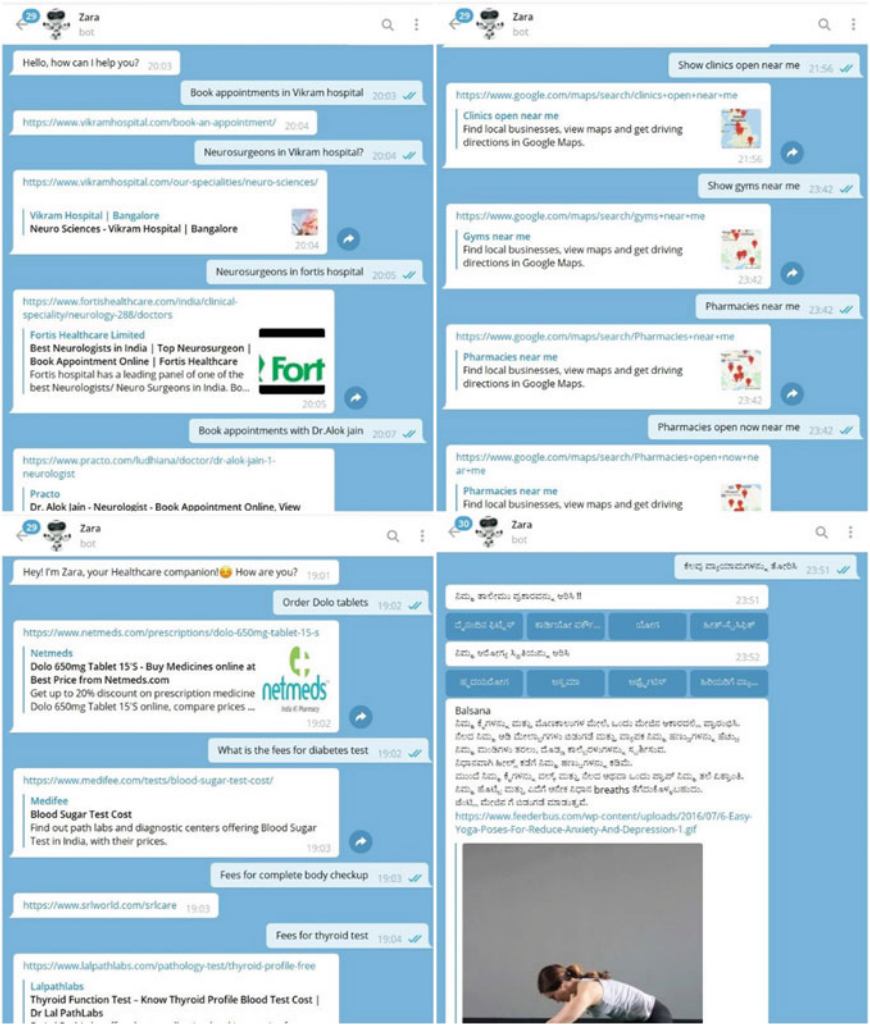



**Fig. 2** Sample conversation with Zara bot

world, India, and also statewise and also informs the users regarding the updates and announcement by the Government of India regarding the same.

## 4 Methodology

This health assistant chatbot is an AI-based or intent-based approach in machine learning. Chatbot is trained with datasets based on natural language processing

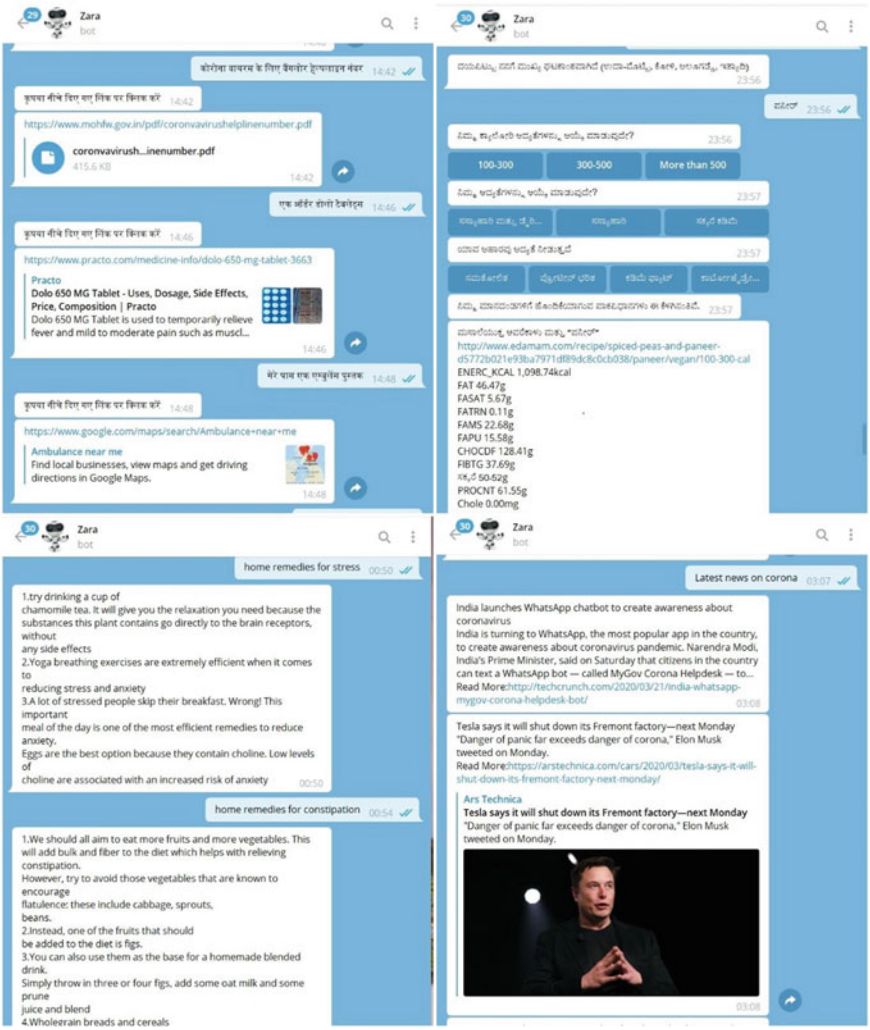



**Fig. 3** Queries addressed by Zara bot

(NLP), which are conversations between the user and the bot for the extraction of the user message. Using an agile development process, efforts were made to improve the bot flow, the conversational knowledge base, and the overall user experience. It was through continuous integration (CI) and continuous development (CD), that a new trained model was developed which could pass all the tests and verification. The testing phase was done side by side with the development phase of the chatbot.

## 5 Functionality

The chatbot is developed using an open-source AI tool and framework Rasa. Rasa mainly includes:

- RasaNLU—a library in Rasa for natural language understanding (NLU) which classifies the intent and extracts the entity from the user input in a conversation and helps Zara to understand what the user is intending to say. Rasa NLU internally uses conditional random field (CRF) to find entities and bag-of-word (BoW) algorithm to find intent.
- Rasa Core—a chatbot framework with machine learning-based dialogue management which takes its structured input from the Rasa NLU. Using this input, training data, and conversation history, the next best activity to be performed by the bot is predicted by employing a probabilistic model.

Zara will be able to understand the user input using the NLU model and training data given to it. The training data consists of core data and NLU data and is stored in the markdown file. Further, the core consists of a file called stories, in which it defines the sample interaction between the user and chatbot in terms of intent and action taken by the bot. NLU data defines the possible user inputs in the form of intents. Also, it stores the entity or keyword which is to be extracted from the user input. Actions are the operations performed by the bot by either integrating with some APIs or querying the database to retrieve or update information. Action file is a Python file which is hosted on a different server.

Rasa also consists of a config file which defines polices and pipelines. This file contains the configurations that will be used in the NLU model. The configuration file is important for the model training as it provides quite a few important parameters which are used while the model is being trained.

Rasa also contains a domain file. This includes the user inputs and actions to be performed in their response and store the necessary information in slots. The domain consists of five key parts consisting of intents, entities, slots, actions, and templates. Each time the data is trained, a model is created in the form of tar file. This can be deployed either on the command line interface or Rasa X. It is a tool set, which can be used to improve the assistant.

## 6 Results

The screenshots of conversations between Zara and user (Table 1).

**Table 1** Test cases

| S . no. | Intents | Utterance | Action |
|---|---|---|---|
| 1 | Show daily routine | Show daily routine exercise | Display the images and instructions by steps to perform exercise |
| 2 | Medicine remind | Remind me to take disprin at 6 pm? | Notify the user at 6 pm with name of the tablet |
| 3 | Recipe ask | Could you show some recipes? | Displaying recipes according to user choices on calories, ingredients, etc. |
| 4 | Ask map | Show hospitals near me | Locate nearby hospitals on map |
| 5 | Remedy | Give me remedies for dandruff | Display the possible remedies |
| 6 | Out of scope | Can I get a hamburger? | Reply that the requests are out of scope |
| 7 | Corona ask | Positive COVID cases in India | Display latest patient count in India |

## 7 Conclusion

Medical chatbots are playing an integral role in the medical industry. Better organization of patient methods, medication management, emergency assistance conditions, or with first aid offer a solution to simple medical problems. This so reduces the burden on medical professionals. Chatbots have ended the quest for digital health by optimizing patient engagement, interaction process and providing medical support to the patients. From the point of view of the user interface, the advancement of chatbots is less concerned with the visual and centers more on a narrative, conversational dimensions. Lastly, a successful implementation of personalized medical assistant can improve the state of health care in the country.

## 8 Future Work

Future scope would include developing an application embedding the existing model with an in-app exercise tracker which will work as a virtual trainer. Image processing can be used for determining the accuracy with which the user performs the exercise. Video submission to the fitness trainers can be included for getting a personalized input to monitor their performance. The bot could also deliver daily newsletters regarding healthy lifestyle to the subscribed e-mail address of the users. The similar features of the bot could be extended to provide support for additional languages.

## References

1. KC GP, Ranjan S, Ankit T, Kumar V (2019) A personalized medical assistant Chatbot. Int J Sci Technol Eng (IJSTE) 5(7)
2. Rashmi D, Neeta AD (2018) A medical ChatBot. Int J Comput Trends Technol
3. Shobha P, Girija H, Harshitha S (2018) Development of data acquisition system based on IoT
4. Hameedullah K, Chowdhry BS, Zeesha M (2016) MedChatBot: an UMLS based Chatbot for medical students. Int J Comput 55(17)
5. Navida Bel- Gaumwala, Rajashekarappa (2019) Chatbot: a virtual medical assistant. Int J Res Appl Sci Eng Technol (IJRASET) 7(6)
6. Raj P, Murali Krishna R, Krishna SM, Vardhan KH (2017) Emergency patient care system using Chatbot. Int J Technol Res
7. Tobias K, Marcia N, Shih CHI, Ruegger D, Volland D, Filler A (2018) Text-based healthcare Chatbots supporting patient and health professional teams

# Detection of Leukemia Using Convolutional Neural Network

**V. Anagha, A. Disha, B. Y. Aishwarya, R. Nikkita, and Vidyadevi G. Biradar**

## 1 Introduction

The disease cancer is the second major cause of death according to Mayo Foundation for Medical Education and Research (MFMER). With increase in world population, the number of deaths due to cancer is also increasing. There is a need for early diagnosis of cancer in order to reduce deaths [1]. The challenge involved is that many people do not depict the symptoms of leukemia in the initial stages. The medical procedure followed for leukemia classification is blood tests which show abnormal White Blood Cell count and aspiration of the bone marrow. These tests need to be analyzed which is time consuming as the doctors need to individually look through each blood sample. This procedure must be improved in cases where timely detection is essential. Therefore, an automatic tool for detection of leukemia is required. Image processing techniques supported by deep learning libraries like Keras may be used to build a Leukemia detection model. This system should be able to overcome the drawbacks of visual investigations and result in the timely identification of the disease.

In [1], authors have considered geometrical features of cell image as distinctive features for classification combined with other statistical parameters like mean and standard deviation to separate white blood cells. Nucleus segmentation is carried out using Otsu's thresholding and Sobel operator and features are extracted. One of the major limitations is that it is not an automated system for classification and the number of features being used for classification is only a few which may hinder the results.

V. Anagha · A. Disha · B. Y. Aishwarya · R. Nikkita · V. G. Biradar (✉)
Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore 560064, India
e-mail: vidyadevi.g.biradar@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_20

A computerized system that detects and classifies acute lymphoblastic leukemia was designed in [2]. Peripheral blood smear (PBS) images are obtained and White Blood Cells are detected by histogram equalization. The triangular algorithm is used for segmentation. Separation of connected cell components is achieved by Watershed segmentation using distance transform image calculation. Finally, Nucleus and Cytoplasm are separated and the features are extracted. A major limitation is that the system needs more testing.

The focus of recent research work for image classification is based on deep learning techniques. Research work in [3] has dealt with the architecture of Convolutional Neural Networks (CNN) and their suitability for image classification. The authors of [4] have designed and implemented CNN model for classification of CIFAR-10 dataset and identified parameters that affect the performance of classification model.

In literature there exist several methods for automating the problem of detecting White Blood Cell cancer using traditional image processing techniques. This work explores deep learning techniques using CNN.

## *1.1 Convolutional Neural Network (CNN)*

Convolutional Neural Network is the most sought after deep learning model for image classification. A CNN model typically consists of three main layers: Input, Hidden, and Fully Connected layer as shown in Fig. 1.

There are two types of transformations available in CNN architecture:

i .Convolution, the pixels in the images are convolved using a filter or kernel which gives the dot product between the kernel and the image patch. Followed by the activation function.

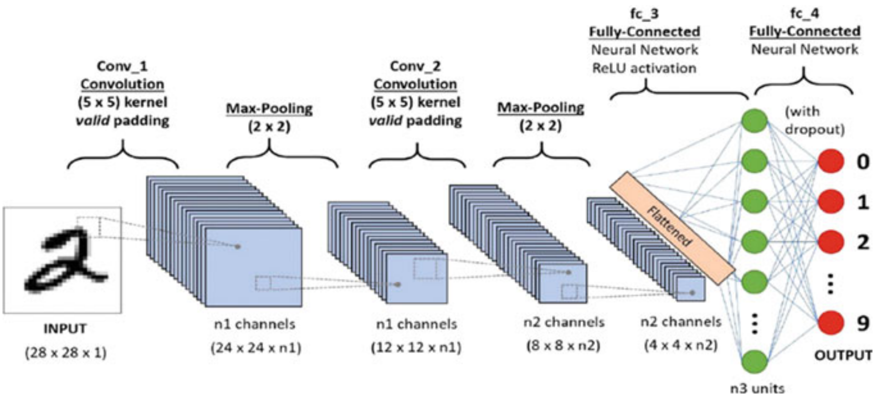



**Fig. 1** CNN architecture

ii .Sub-sampling is the other important transformation which can be of three types namely, max, min, and average pooling. This layer reduces the dimensionality of the data using a pooling filter whose size is set by the user.

After a combination of these two layers, the output is passed to the fully connected layer for classification.

The focus of the proposed system is the exploration of deep learning techniques using Convolutional Neural Network (CNN) to solve the problem of Leukemia classification. Deep learning techniques have become very popular due to their high performance in image classification problems as compared to traditional image processing methods [2].

The paper is organized into different sections: Section 2 discusses various approaches for Leukemia detection, Sect. 3 describes the methodology of the proposed model, Sect. 4 deals with the results and discussion and Sect. 5 contains conclusion and future work.

## 2 Related Work

In [5], authors have designed a system that classifies the dataset of 100 images into the 6 different subtypes of leukemia. Input images are converted into YCBCR color space, segmented using Gaussian Distribution, and feature extraction is done using GLCM. The system obtained an accuracy of 97.3%.

The authors of [6] mainly focus on nucleus segmentation and feature extraction. The preprocessing steps filtering and image enhancement using histogram equalization Segmentation are carried out based on mean and standard deviation. Ostu's thresholding and filtering are used to enhance the nucleus to reduce classification errors. The paper concludes that the shape features of the nucleus have to be considered to achieve better accuracy of detection.

The various methods that can be used for detection of leukemia are discussed in [7]. They include Watershed transform, *K*-means clustering, HIS color model based classification, and model based on shape features. The study indicates the model based on shape-based features is the most accurate to detect the different types of cells present in the image.

An analysis of the various techniques of image processing used in the identification of leukemia is given by [8]. Digital data of microscopic images of samples are collected and the white blood cells are segmented. K Means method along with Fuzzy C-Means method was found to have a relatively high degree of accuracy. It was concluded that carrying out feature selection beforehand through T-SNE, LDA or PCA was a better approach. Based on the individuality of cells and the characteristics of their morphological structures, cytoplasm, and nucleus, cells are differentiated as cancerous and non-cancerous cells. It was found that SVM and MLP methods produce accurate classification results.

In [9], the authors have discussed the information technology used in the detection and prevention of leukemia. Linear Dependent Analysis is used as a dimensionality reducing method during preprocessing. Artificial Neural Network (ANN) mimics the exact behavior of the brain by means of neural link which can be used in differentiating the affected cells. Self-Organizing Map is an unsupervised-learning technique used for feature extraction and classification of cells is done using SVM. Genetic Algorithm is used to derive the system behavior. This includes the formation of population followed by evaluation of its fitness, selection, crossover, and mutation. This paper explains the different techniques used for leukemia cell detection.

Rejintal and Aswini [10] proposes a system with two parts-testing and training. First, image acquisition with proper magnification from any hospital is done. Then, image preprocessing where color images are converted into gray scale, noise is removed to increase quality. Segmentation is done using K means clustering and nucleus is considered for detection. Features like correlation, entropy, and contrast are extracted by using methods like GLCM and GLDM. During training, the features of the cancer cells are saved in the knowledge base and during testing, images to be tested are taken as input. SVM classifier is utilized to classify the images as having cancer or not. It is concluded that k means method is best suited for segmentation as the performance is high.

In [11], the authors have designed a system that aims to detect leukemia and determine its types. The 220 blood smear images were converted to RGB form and segmentation algorithms-HSV color-based segmentation, Marker controlled Watershed and K means clustering was applied. Features like cell-size, mean, entropy, standard deviation, correlation, variance are extracted and an SVM classifier with the dataset split into training and testing is used to classify the cells. This paper concludes that the features extracted play an important role in the detection of leukemia.

Image preprocessing and classification into the subtypes of leukemia is the main focus of [12]. The images are first converted into CIE l, a, b space from the RGB color space to differentiate between the cytoplasm and the nucleus. Cells maybe are overlapped in some images which are identified using linear interpolation and separation is done by getting the midpoint of all the curved portions. Wold's decomposition is used to determine the texture of the image from the data set. Features like area, correlation, entropy, mean are used for classification. SVM and ANN fused classifier increases performance.

Jasmine Begum and Razak [13] discusses the complete process of detection of leukemia. The proposed model converts the microscopic RGB image into Gray scale image, enhances the contrast, applies Otsu's Global threshold method for segmentation, reduces noise, preserves edges, and increases the darkness of nuclei using Morphological Min filter resulting in images of affected cells. Features like area, perimeter, and circularity are extracted and eccentricity, solidity is calculated. Support vector machine is used for classification to find the best hyperplane that represents the largest margin between two classes. It concludes that the shape of the nucleus plays a major role in the accurate detection of the leukemia blast as compared to the shape of the cytoplasm.

A system to identify the types of leukemia namely, Acute Lymphoblastic Leukemia and Acute Myelogenous Leukemia is designed in [14]. Three preprocessing steps, Median Filtering to repair the White blood cells, RGB to HSV conversion, Thresholding to separate the WBC from the background, are applied to the original images obtained from the microscopic. Watershed and Integral Projection methods are used for segmentation and feature extraction respectively. Identification was done using Standard deviation and if-else branching. This system gave high accuracy results.

Mohammed et al. [15] categorizes the disease into two based on similar visual features and two approaches were implemented where different features are computed. One differentiates between M5 AML, L1, and L2 ALL and the other differentiates the remaining subcategories. During preprocessing, images were converted from RGB to YCbCr color space, the nucleus and cytoplasm were segmented using the Cb and Cr co-efficient by building the Gaussian Distribution. Morphological features like circularity, statistical features like standard deviation, size ratio features like nucleus cytoplasm area, and texture features are extracted. Random Forest algorithm is used for classification with highest accuracy. The system was capable of relearning from the misdiagnosed cases.

## 3 Methodology of Proposed Model

The research literature shows that there is a scope for development of a novel classifier for Leukemia detection. In this work, a Leukemia classifier model is implemented using Convolutional Neural Network.

Figure 2, shows a sequence of steps followed to build a prediction model using CNN.

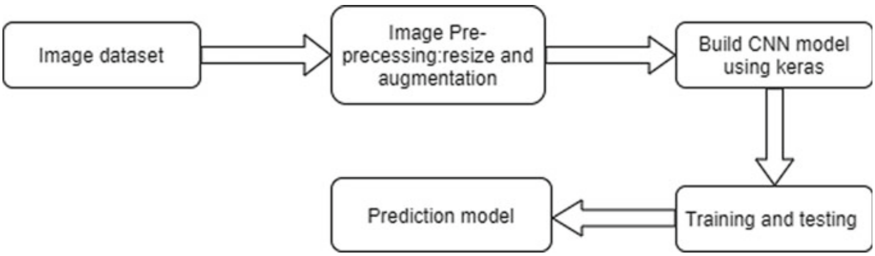



**Fig. 2** Steps in development of Leukemia prediction model

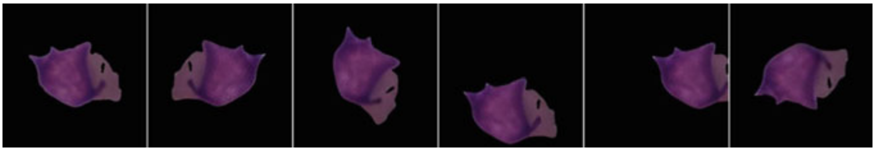


**Fig. 3** Example of augmented dataset of cell images

### 3.1 Dataset

The dataset consists of 2343 White Blood Cells that have been segmented from microscopic images of blood smears. The cell images are of 2 types, Healthy and Acute Lymphoblastic Leukemia: 1173 ALL and 1170 healthy. The images contained in the dataset are in BMP format with a size of 450 × 450. The dataset was taken from The Cancer Imaging Archive (TCIA) Public Access [16].

### 3.2 Tools Used

Deep learning library Keras is used with TensorFlow as back end. The visualization of feature extraction of White Blood Cells at intermediate layers of CNN is implemented using Lime tool.

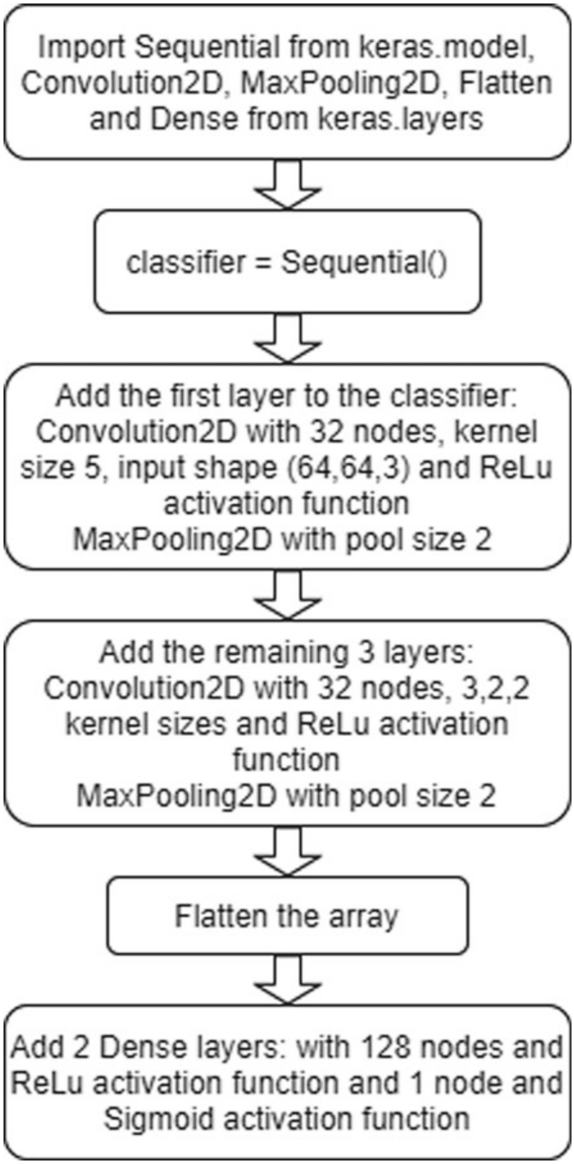
### 3.3 Image Preprocessing

Image preprocessing includes the resizing of images and image augmentation. Images are resized to 64 × 64. Image augmentation is carried out to increase the number of samples in the cancer cell dataset which is an essential step when dealing with medical images as obtaining authenticated medical images is challenging [17–20]. Image augmentation also avoids model over-fitting as the performance of the deep learning models increases with an increase in dataset. Image augmentation is carried out by using methods like rotation, width shift, height shift, horizontal flip, and vertical flip. Figure 3 shows examples of augmented images.

### 3.4 Architecture of Proposed CNN Model

1 The model was built using a sequential model as shown in Fig. 4.
2 Four hidden layers were added, each one consisting of a convolutional layer followed by a pooling layer.

**Fig. 4** CNN Model




3 Parameters chosen for convolutional layers are, filters-32, kernel size-5 $\times$ 5, ReLu (Rectified Linear Unit) activation function for thresholding.
4 Pooling layer filter size chosen is 2 $\times$ 2 and max pooling operation is applied.
5 The 2D features obtained from last convolutional layer are converted into 1D feature vector by making use of a flattened layer.

| Layer (type) | Output Shape | Param # |
|---|---|---|
| conv2d_1 (Conv2D) | (None, 60, 62, 32) | 1472 |
| max_pooling2d_1 (MaxPooling2 | (None, 30, 31, 32) | 0 |
| conv2d_2 (Conv2D) | (None, 28, 29, 32) | 9248 |
| max_pooling2d_2 (MaxPooling2 | (None, 14, 14, 32) | 0 |
| conv2d_3 (Conv2D) | (None, 13, 13, 32) | 4128 |
| max_pooling2d_3 (MaxPooling2 | (None, 6, 6, 32) | 0 |
| conv2d_4 (Conv2D) | (None, 5, 5, 32) | 4128 |
| max_pooling2d_4 (MaxPooling2 | (None, 2, 2, 32) | 0 |
| flatten_1 (Flatten) | (None, 128) | 0 |
| dense_1 (Dense) | (None, 128) | 16512 |
| dense_2 (Dense) | (None, 1) | 129 |

Total params: 35,617
Trainable params: 35,617
Non-trainable params: 0

**Fig. 5** CNN model summary

6 A dense layer having 128 nodes with ReLu activation function follows the flatten layer. The fully connected layer or output layer consists of one node with Sigmoid activation function.

The CNN model is visualized with its summary as given in Fig. 5.

## 3.5 *Training Phase*

The learning configuration for CNN is set with the parameters, adam optimizer, the loss function is set as binary cross entropy, and the metric used to train the model is accurate.

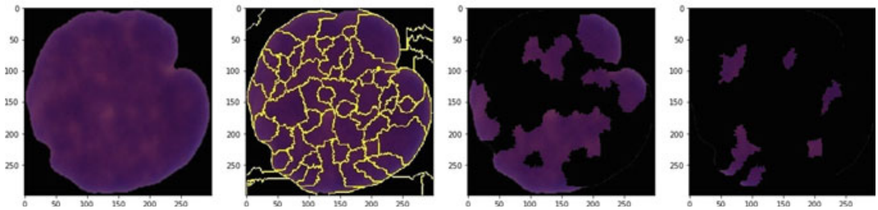


**Fig. 6** Features of acute lymphoblastic Leukemia cell

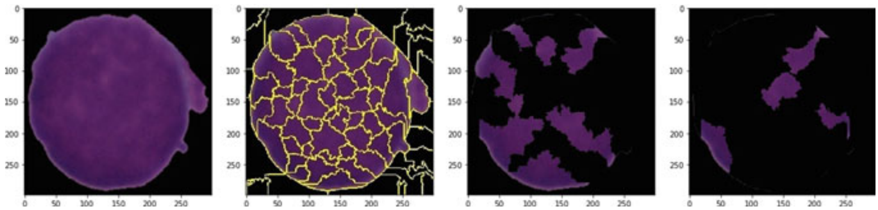


**Fig. 7** Features of healthy cell

The dataset is split into training set and testing set. The size of the testing set to 25% of the dataset which has 586 images and the remaining 1757 images are used for training. The model trains on the training set and learns the features that distinguish between the two classes of cells (Acute Lymphoblastic Leukemia and healthy). Training is carried out by varying the number of epochs.

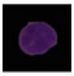
There are two possible predictions made by the model. One is Acute Lymphoblastic Leukemia and the other is Non-leukemic or Healthy. When an image is given as input to the model, it passes through multiple layers of convolution and pooling and it learns specific features in each layer which helps in making the final class prediction. The Figs. 6 and 7 shows some of the features based on which the model predicts the class of the cells.

## 4 Testing

The model was tested on 20 new images and outcomes are tabulated as shown in Table 1.

As seen from Table 1, the classifier provides good predictions for each category.

**Table 1** Test results for 20 new images

| S. No | Image | Actual class | Predicted class | Result |
|---|---|---|---|---|
| 1 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 2 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 3 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 4 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 5 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 6 |  | Acute Lymphoblastic Leukemia | Healthy | False |
| 7 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 8 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 9 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 10 |  | Acute Lymphoblastic Leukemia | Acute Lymphoblastic Leukemia | True |
| 11 |  | Healthy | Healthy | True |
| 12 |  | Healthy | Healthy | True |
| 13 |  | Healthy | Healthy | True |

(continued)

**Table 1** (continued)

| S. No | Image | Actual class | Predicted class | Result |
|---|---|---|---|---|
| 14 |  | Healthy | Healthy | True |
| 15 |  | Healthy | Healthy | True |
| 16 |  | Healthy | Acute Lymphoblastic Leukemia | False |
| 17 |  | Healthy | Healthy | True |
| 18 |  | Healthy | Healthy | True |
| 19 |  | Healthy | Healthy | True |
| 20 |  | Healthy | Healthy | True |

## 5 Results and Discussion

The dataset containing 2343 images was split into training and testing with a test train split of 0.25% and 75%. The proposed model gives an accuracy of 91% and 87% for training and testing sets respectively. The Fig. 8 shows the accuracy vs epoch graph of the system. It is evident that the accuracy achieved by the system is increasing with each epoch. The Fig. 9 depicts the loss vs epoch graph of the model for 180 epochs. The loss is decreasing with each passing epoch. Lower values of loss indicate that the model's performance is better. As the accuracy of the system increases, loss of the system decreases, and thus the probability of the system making an error is less.

The model's performance was evaluated for precision, recall, F1-score, and support score metrics of the model. The model has an F1-score of 0.87 and a precision of 0.89 for class Acute Lymphoblastic Leukemia and 0.86 for class Healthy. The results are given in Fig. 10, indicate that the performance of the proposed model is satisfactory.

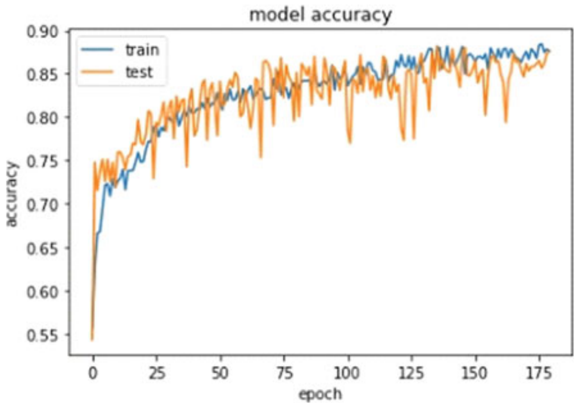



**Fig. 8** Accuracy versus epochs

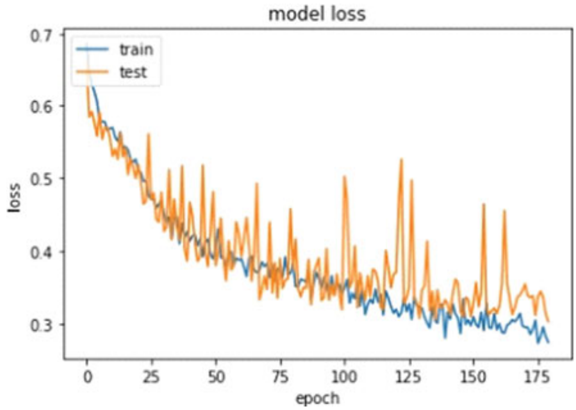



**Fig. 9** Loss versus epochs

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.89 | 0.86 | 0.87 | 298 |
| 1 | 0.86 | 0.89 | 0.87 | 288 |
| accuracy | | | 0.87 | 586 |
| macro avg | 0.87 | 0.87 | 0.87 | 586 |
| weighted avg | 0.87 | 0.87 | 0.87 | 586 |

**Fig. 10** Model performance metrics

# 6 Conclusion

This paper presents CNN based prediction model for detection and classification of white blood cell cancer. Performance of the model is evaluated on C_NMC_2019 dataset which contains white blood cell regions segmented from the microscopic blood smear images. The model was developed using deep learning library Keras with backend as TensorFlow. The model gives an accuracy of 91% for the training set and 87% for the test set. The model also has an F1-score of 87%. Hence, the proposed model gives good predictions. The future work is to explore the suitability of pre-trained deep learning models with transfer learning techniques, also Capsule Networks may be explored as they incorporate neighborhood information.

## References


1. Khobragade S, Mor DD, Patil CY (2015) Detection of Leukemia in microscopic white blood cell images. In: International conference on information processing(ICIP)
2. Rahman A, Hasan MM (2018) Automatic detection of white blood cells from microscopic images of malignancy classification of acute lymphoblastic Leukemia. In: 2018 international conference on innovation in engineering and technology(ICIET)
3. Albawi S, Mohammed TA, Saad AZ (2017) Underrstanding of a convolutional neural network. Int Conf Eng Technol (ICET)
4. Chauhan R, Ghanshala KK, Joshi RC (2018) Conventional neural network(CNN) for image detection and recognition. In: 1st international conference on secure cyber computing and communication(ICSCCC)
5. Agrawal R, Satapathy S, Bagla G, KR (2018) Detection of white blood cell cancer using image processing. In: International conference on vision towards emerging trends in communication and networking (ViTECoN)
6. Raje C, Rangole J (2014) Detection of Leukemia in Microscopic Images Using Image Processing. In: International conference on communication and signal processing
7. Veghela HP, Modi H, Pandya M, Potdar MB (2015) Leukemia detection using digital image processing techniques. Int J Appl Inf Syst (IJAIS)
8. Bagasjvara RG, Candradewi I, Sri H, Harjoko A (2016) Automated detection and classification techniques of acute leukemia using image processing: a review. In: 2nd international conference on science and technology-computer (ICST)
9. Singh G, Bathal G, Kaur SP (2016) A review to detect Leukemia cancer in medical images. In: International conference on computing, communication and automation (ICCCA)
10. Rejintal A, Aswini N (2016) Images processing based Leukemia cancer cell detection. In: IEEE international conference on recent trends in electronics, information and communication technology(RTEICT)
11. Jagadev P, Virani HG (2017) Detection of Leukemia and its types using processing and machine learning. In: International conference on trends in electronics and information(ICEI)
12. Shaikh MB, Deshpande S (2017) Computer aided leukemia detection using digital images processing techniques. In: 2nd IEEE international conference on recent trends in electronics, information and communication technology(RTEICT)
13. Jasmine Begum AR, Razak TA (2017) Diagnosing Leukemia from microscopic images using image analysis and processing techniques. In: World congress on computing and communication technologies(WCCCT)
14. Sigit R, Bachtiarand MM, Fikri MI (2018) Identification of Leukemia diseases based on microscopic human blood cells using image processing. In: International conference on applied engineering (ICAE)
15. Mohammed H, Omar R, Saeed N, Essam A, Ayman N, Mohiy T et al (2018) Automated detection of white blood cells cancer diseases. In: 1st International workshop on deep and representation learning(IWDRL)
16. (n.d.) Retrievedfrom https://wiki.cancerimagingarchive.net/display/Public/C_NMC_2019+Dataset%3A+ALL+Challenge+dataset+of+ISBI+2019

17. Vasconcelos CN, Vasconcelos BN (2017) Incresing deep learning melanoma classification by classical and expert knowledge based image transforms
18. Zhou J, Li Z, Zhi W, Liang B, Moses D, Dawes L (2017) Using convolutional neural networks and transfer learning for bone age classification. In: IEEE international conference oon digital image computing techniques and applications(DICTA)
19. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H (2018) Gan Based synthetic medical image augmentation for increased CNN performance in liver lession classification
20. Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J et al (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning

# TORA: Text Summarization Using Optical Character Recognition and Attention Neural Networks

**H. R. Sneha and B. Annappa**

## 1 Introduction

We live in a generation where enormous amounts of data are generated every single day. In fact, the International Data Corporation (IDC) projects that the total amount of digital data circulating annually around the world would sprout from 4.4 zettabytes in 2013 to hit 180 zettabytes in 2025 [1]. Ironically, the amount of time people can dedicate to read all this data is very less. Thus, there is a necessity to reduce this Large amount of text data to more concise summaries that capture the salient details of the original data. Thus, comes the concept of text summarization.

Text Summarization is the process which helps in reducing the length of the data and create summaries that hold the same meaning as that of the original data [2]. Text summarization reduces reading time and expedites the process of researching for information.

Automatic Text Summarization [3] is an extension of the concept of Text Summarization. Automatic Text Summarization is the process of creating a short and coherent version of a longer document using machine learning techniques. The aim of automatic text summarization is to create summaries that are as good as summaries that have been created manually. It is not sufficient that the algorithm generates only relevant words and phrases from the original text, rather it should be able to generate sentences that are grammatically correct.

H. R. Sneha (✉)
Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru 560064, India

B. Annappa
Department of Computer Science and Engineering, National Institute of Technology Karnataka, Surathkal 575025, India
e-mail: annappa@ieee.org



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_21

Text Summarization is not restricted to news articles but it may also be used for various other applications such as abridgment of books or novels, abstract of a scientific paper, minutes of a meeting, outline of notes for students, etc. This research work is designed to create summaries that are concise and at the same time convey the same message as that of the original text. The text summarization method proposed in this article will be validated in newspaper articles. The goal and objectives of the paper are,

- Develop methods for reading the input from various means—one from digital textual document, and other from scanning the physical document using Optical Character Recognition to read the content for text summarization.
- Preprocessing the input data to remove unwanted characters.
- Apply Recurrent Neural Networks, Long Short-Term Memory, and Attention Networks to build the text summarization models using tensor flow framework.
- Validate the generated (Tensor flow) model on the data read using the method mentioned in objective one (digital document and physical document).

The upcoming sections of this paper are organized as follows. Section 2 briefs the related works in on the context of the paper, Sect. 3 gives the different methods for text summarization, Sect. 4 gives the different machine learning models used for text summarization. Section 5 give the mathematical modeling and framework of the project. Section 6 discusses results of the proposed method on newspaper articles taken from internet. The final part of the paper summarizes the contribution of the paper as a brief conclusion.

## 2 Related Work

Automatic Text Summarization can be classified on the basis of either the input type or purpose or the output type. Input Type can be Single or Multiple Document, Purpose can be of Generic Domain, Specific or Query Based. Output types are Extractive or Abstractive [4]. The overview of subfields in text summarization is shown in Fig. 1.

### *2.1 Text Summarization Based on Output*

We focus on the Automatic Text Summarization classification based on the output type, we propose to demonstrate one of the output based approaches classified as Extractive Text Summarization [2], and Abstractive Text Summarization [5].

Extractive Text Summarization is the process of picking sentences directly from the original text rather than creating the sentences from scratch. It is important that the correct important sentences are identified by the algorithm. All the sentences that are identified are later concatenated to form the summary.

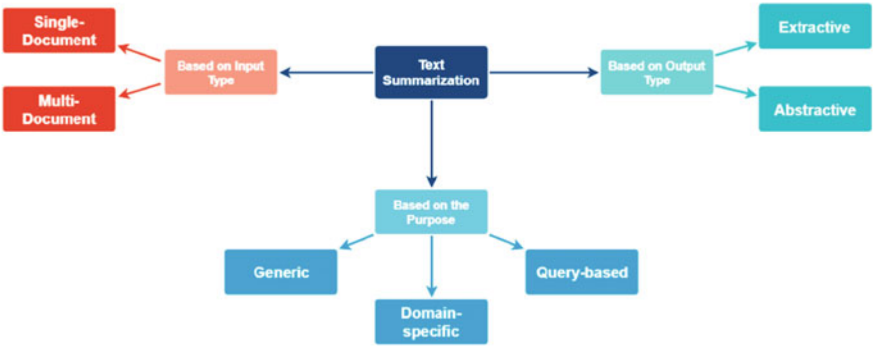



**Fig. 1** Types of text summarization

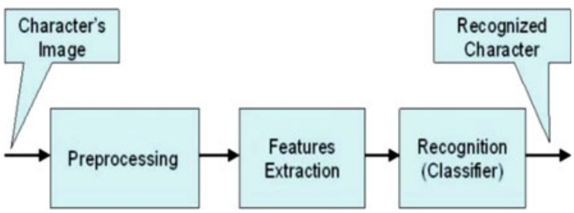
### *2.2 Sentence Extraction Methods*

The three main approaches to sentence extraction are: Frequency based approach, Feature based approach, and Machine learning approach [6]. Abstractive text summarization is a text summarization technique that generates new sentences altogether rather than selecting the ones already present in the document. The summaries generated might contain words and phrases that might not be present in the original document. Abstractive text summarization uses an approach similar to that used by humans and thus it is a more difficult approach to summarizing a document. Abstractive methods can be classified as follows structure based approach, and semantic based approach.

## 3 Optical Character Recognition Using Python-Tesseract and One Hot Encoding

Optical Character Recognition (OCR) is a tool that recognizes printed or written text characters with the help of a computer. It involves photo scanning of the text character by character, analysis of the scanned image, and the translation of the character image into ASCII codes. Overview of working of OCR is given in Fig. 2.

**Fig. 2** Working of optical character recognition



Firstly, we need a document in textual format to extract the words and create the summary. Suppose we have an image of the news article; we need to convert the data given in the image into a text format. This is where we use OCRs to identify the characters and store them as digitized text documents.

### 3.1 OCR Using PyTesseract

Python-tesseract (PyTesseract) is an Optical Character Recognition (OCR) tool for python. It extracts the text embedded in images. It is a wrapper for Google's Tesseract-OCR Engine. It is open source and is developed and maintained by Google. It is also useful as a stand-alone invocation script to tesseract, as it can read all image types supported by the Python Imaging Library, including jpeg, png, gif, bmp, tiff, and others, whereas tesseract-OCR by default only supports tiff and bmp. Additionally, if used as a script, Python-tesseract will print the recognized text instead of writing it to a file.

### 3.2 One Hot Encoding

Machine learning algorithms cannot work directly with categorical data and hence it must be converted to numbers when working with classification problems or deep learning methods.

One hot encoding involves the mapping of the words to a vector form. The entire word list from the dataset is acquired in the previous step. One hot encoding is a representation of categorical variables to integer values. Each integer value is represented as a binary vector that is all zero values except the index of the integer that is represented with one. Mathematically, one hot encoding produces a balanced matrix, which is easy to understand during complex computations inside algorithms. One such example is shown in Fig. 3, where each city is encoded as a vector consisting of a set of zeros and ones.

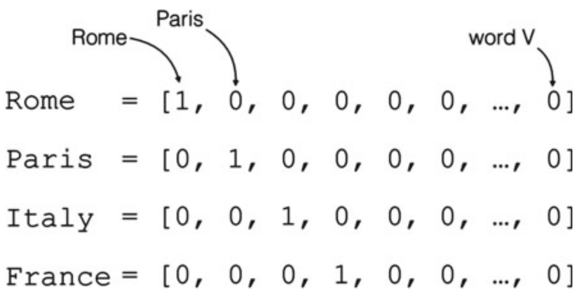



**Fig. 3** One hot encoding for categorical to numerical conversion

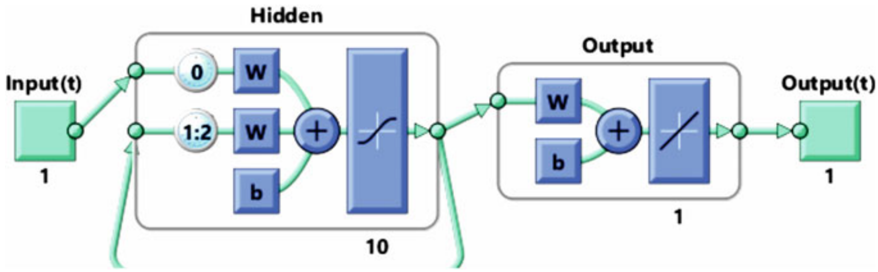



**Fig. 4** Recurrent neural networks

# 4 Experimental Setup

## *4.1 Recurrent Neural Network, Long Short-Term Memory and Attention Networks*

Recurrent Neural Networks are a type of neural network that have the ability to retain information from its previous inputs and carry it to the next stage along with the new input (Fig. 4).

$$h(t) = f\left(h^{(t-1)}, x^{(t)}; \theta\right)$$

This formula explains that the current hidden state $h(t)$ of any network is a function $f$ of the previous hidden state $h(t-1)$ and the current input $x(t)$. The theta denotes the other parameters of the function $f$. Activation functions are used to determine the output of neural in a yes or no format. It maps the result calculated by the neural network in values that lie between 0 to 1 or $-1$ to 1 which depends on the type of function used.

## *4.2 Activation Functions*

Various linear and Nonlinear activation functions are used in various layers of attention networks constructed for text summarization. On the high level, activation functions function is classified as linear and nonlinear. Output of Linear activation function is not confined between any range. It does not help with the complexity of the parameters of the usual data that is used as an input to neural networks. On the other hand, nonlinear activation functions have a confined range—these activation functions are the most commonly used activation functions for constructing neural networks. The different types of activation functions are softmax, sigmoid, tanh, ReLU, and Leaky ReLU. In this work, we use softmax, sigmoid, and tanh activation functions.

The softmax function (shown in Fig. 5) normalizes an input value into a vector of values that follows a probability distribution whose total sums up to 1. Sigmoid or Logistic Activation Function is an S-shaped function. This function is confined between the range 0–1. It is used for models where the output to be predicted is in the form of a probability. This function can be differentiated to find the slope of the sigmoid curve at any two points. Tanh Function is also called the hyperbolic tangent activation function. The range for this type of function is confined between −1 to 1. It is also an s-shaped function (Fig. 6).

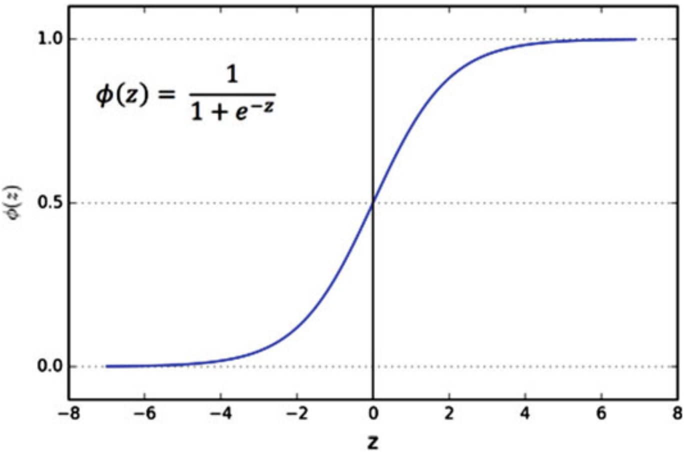
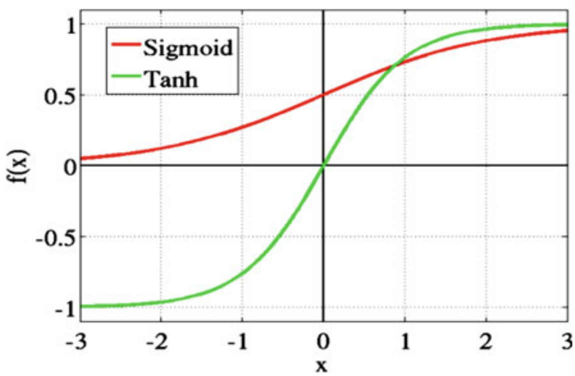



**Fig. 5** Softmax activation function

**Fig. 6** Sigmoid and Tanh activation function



### 4.3 Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) Networks (shown in Fig. 7) are capable of handling the long-term dependencies as they can remember information for a long period of time. The model that we are using in this work is a modified version of the recurrent neural networks. It has the capacity of an internal memory that can last for a very long time.

RNNs can connect the previous information to the present task but at times, we need more context or information from the past. This connection from a relevant point to the current task may be very large and cannot be covered up using RNNs. These are called "long-term dependencies". Thus, here we used a modified version of RNNs to help understand the next word to be predicted. These long-term dependencies are covered by what is called Long Short-Term Memory (LSTM). In this work, we use these LSTM's to predict the new word for the summary.

Every RNN layer will have a tanh module to bind the previous output with the present input at every stage. But, in case of LSTM's, this chain like structure also has a repeating module having four segments with different functionalities at each stage.

In LSTM network model shown in Fig. 7, each line is a vector that carries the result or solution from one node to another node. The pink circles denote every operation module. The yellow boxes are the gates that define a way to optionally let information through. They are composed of either the sigmoid function or the tanh function.

Forget gate layer decides the information to be forgotten or thrown away from the memory. This is achieved by a sigmoid function of the previous layer and the current inputs to output a value between 0 and 1 where 0 represents the forget state and 1 represents the keep state.

Store layer decides what new information is to be stored. This is achieved with 2 parts. First input gate layer sigmoid function layer to decide which values to update.

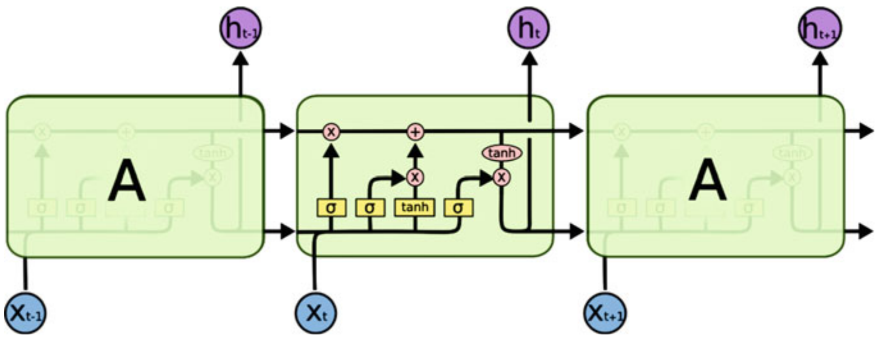



**Fig. 7** Long short-term memory (LSTM) networks

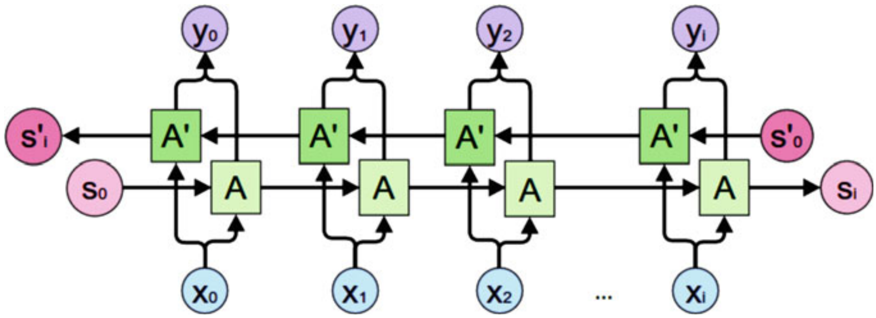



**Fig. 8** Bidirectional long short-term memory (Bi-LSTM) networks

Secondly, tanh layer is used to create a vector of new values that could be added to the state or to the memory.

### *4.4 Bidirectional Long Short-Term Memory*

The limitations of RNN that it can only predict the forward states only in a positive time direction can be overcome by modifying the LSTM network to involve both forward and backward states. This is used when the sequence generation of a sentence is not only dependent on the words that are present before it but also on the words that are present after it. This way it is required to know the positive and negative directions of the context and Bidirectional LSTM's play a major role to calculate (Shown in Fig. 8).

### *4.5 Attention Networks*

The modified RNN's, i.e., Attention Network shown in Fig. 9, work well with sequence generation programs but several new properties are now added to the recurrent neural network and are called augmented recurrent neural networks. Attention Networks is a mechanism that is used to analyze a sequence in a similar manner to that of a human interpretation of translating a sentence which is mainly done part by part in any sentence. This mechanism has proven to provide more accuracy in machine translation and natural language processing.

A RNN crams the knowledge about the data elements from every sequence into the last hidden state of the network while an attention mechanism takes into account the input from several previous hidden states and uses them to make one prediction. It allocates different weights and degrees of importance to the different hidden layers of

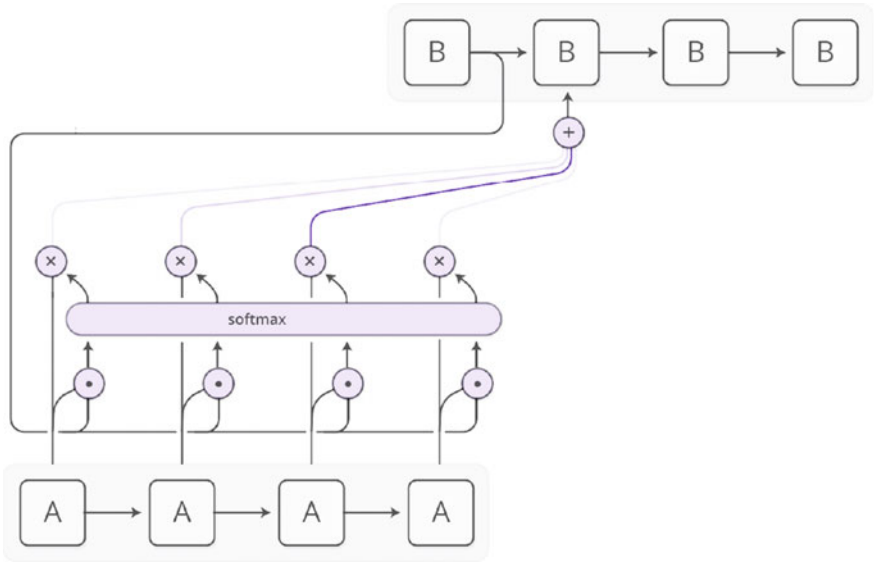



**Fig. 9** Attention network

the neural network. Attention acts as a memory-access mechanism that is responsible for identifying which areas of the sentences or the word are important in the sentence.

## 5 Results and Discussion

The dataset that we have used in the model is taken from two different sources: Dataset from Kaggle that in news domain that contains 143,000 from 15 American publications like CNN, New York Times, Business Insider, etc. It falls between the years 2016 and 2017. Dataset from Times Of India contains nearly 1300 articles which fall between the years of 2018–2019.

### *5.1 Dataset Preprocessing*

The dataset might have many discrepancies such as missing data fields, html tags, or escape sequences. We have incorporated functions that perform the following:

- Any row in the dataset that has a missing headline or story is dropped so there are no mismatches.
- Remove html tags, escape sequences, and punctutations.
- Convert the paragraphs into sentences and then words.

- Add the words to the word dictionary.

### 5.2 Training the Model

The dataset obtained after preprocessing is used to train the model. A part of the dataset is used to train the model. It is a combination of articles from both of the data sets. It consists of articles from publishers like BBC News, New York Times as well articles published by Times of India. The length of these articles ranges from 50 characters upto a maximum of 1000 characters per article. In this project, the training dataset contains about 7000 articles. Before the model is trained, the model parameters need to be set. The following values have been set for the model. The model is allowed to train for 100 epochs with the condition that the training will stop if there is no improvement in the training loss for 5 consecutive epochs. The model parameters are briefed in Figs. 10 and 11.

The model is trained in batch sizes of 32 each. For each interaction in an epoch, the training loss is calculated. At the end of all the interactions in an epoch, the average training loss for that epoch is also calculated. On average, the model gets trained for 84 epochs. Once the training is complete, the new article can be given to the model to predict the summary. The diagram below shows the ow of the project right from the training to the prediction.

The trained model can now be given new input. This input may be in the form of text or an image. The input is given through the front end. The input is passed to the model that predicts the summary and that summary is displayed on the screen. The summarizer can predict the summary accurately to upto 80%. There is a lot of scope for improvement. Training the model with a Larger dataset will improve its accuracy further. Below are screenshots of some of the predictions that have been made by the summarizer.

```
# build graph and train the model
summarizer_model_utils.reset_graph()
summarizer = Summarizer.Summarizer(word2ind,
                                   ind2word,
                                   save_path='./models/headlines/my_model',
                                   mode='TRAIN',
                                   num_layers_encoder = num_layers_encoder,
                                   num_layers_decoder = num_layers_decoder,
                                   rnn_size_encoder = rnn_size_encoder,
                                   rnn_size_decoder = rnn_size_decoder,
                                   batch_size = batch_size,
                                   clip = clip,
                                   keep_probability = keep_probability,
                                   learning_rate = learning_rate,
                                   max_lr=max_lr,
                                   learning_rate_decay_steps = learning_rate_decay_steps,
                                   learning_rate_decay = learning_rate_decay,
                                   epochs = epochs,
                                   pretrained_embeddings_path = pretrained_embeddings_path,
                                   use_cyclic_lr = use_cyclic_lr,
                                   summary_dir = summary_dir)
```

**Fig. 10** Model parameters

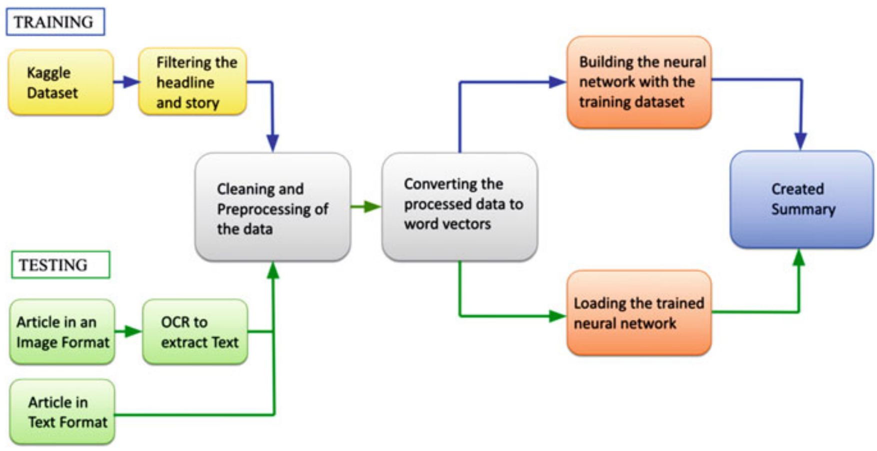



**Fig. 11** Model parameters

Actual text

MUMBAI: The Bombay high court on Thursday frowned at CM Devendra Fadnavis, saying although he holds the home portfolio, he finds no time to take stock of the case related to the killing of CPI leader Govind Pansare. "What is the CM doing? He holds 11 portfolios including home department but does (sic) not find the time to take stock of the case and remove obstacles and hurdles," remarked a bench of Justices Satyaranjan Dharmadhikari and Burgess Colabawalla, which is monitoring the probe. On March 14, the court had said the attempt of the State CID's Special Investigation Team to trace the absconding accused has been reduced to a "laughing stock." Counsel for the state, senior advocate Ashok Mundargi said the strength of the investigation team has been increased to nearly 35.

Actual summary: hc pulls up Maharastra cm over panshare probe

Generated summery: hc pulls ways Maharastra for panshare (Figs. 12 and 13).

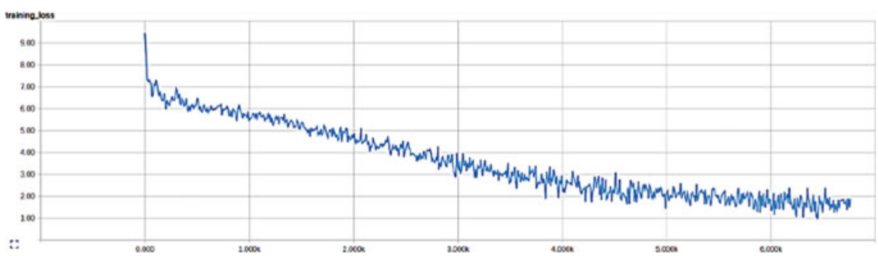



**Fig. 12** Training loss graph

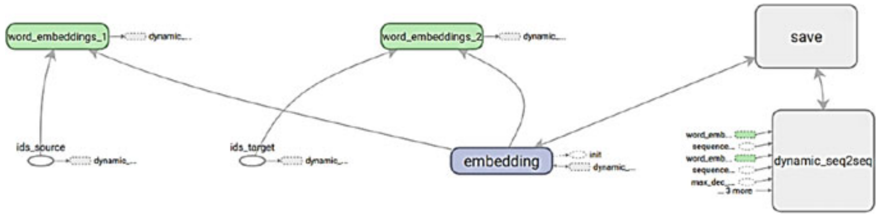



**Fig. 13** Training model graph

NEW DELHI: The tax department has said it is premature to estimate any shortfall in direct tax receipts for the current financial year, which ends in March, and said the actual picture would be known by the middle of April. Reports have suggested that the Central Board of Direct Taxes (CBDT) has shot off a letter to field formations to go all out to bridge the shortfall in direct taxes. The reports have said that direct tax receipts may fall short by 15% of the target.

"This is a routine letter we send out to field formations in the third quarter of the financial year," said an official, who did not wish to be identified. "How can we guess what the shortfall is? Taxes keep coming till April 1 and residuary taxes until April 15," the official said, adding that the final picture will emerge early next month. Earlier this year, CBDT had expressed confidence about meeting the target of Rs. 11.5 lakh crore for the current fiscal year.

Actual summary: Premature to estimate direct tax shortfall

Generated summary: Premature to estimate direct tax shortfall times.

## 6 Conclusion

This work attempts at summarizing news articles of different domains. Just as the in shorts app, which summarizes news articles in 60 words with manual human summarization. This project can also be extended and modified for summarization in other domains such as science, law, and corporate sector for summarizing scientific papers, judicial documents, and minutes of a meeting. The proposed model is cost effective as it only requires initial development cost. It is also a quicker alternative to summarization by humans as it saves time and effort. In the future, the model can be improvised by training it with a Larger dataset. It can be trained with news articles from publishers all over the world to make it a universal summarizer.

## References

1. Ramgovind S, Eloff MM, Smith E (2010) The management of security in cloud computing. In: Information security for South Africa, IEEE, pp 1–7
2. Gupta V, Lehal GS (2010) A survey of text summarization extractive techniques. J Emerg Technol Web Intell 2(3):258–268
3. Tas O, Kiyani F (2007) A survey automatic text summarization. Press Acad Procedia 5(1):205–213
4. Hovy E, Lin C-Y (1998) Automated text summarization and the SUMMARIST system. In: Proceedings of a workshop on held at Baltimore, Maryland. Association for Computational Linguistics, pp 197–214
5. Nallapati R, Zhou B, Gulcehre C, Xiang B (2016) Abstractive text summarization using sequence-to-sequence rnns and beyond. arXiv:1602.06023
6. Ravikumar S, Ramachandran KI, Sugumaran V (2011) Machine learning approach for automated visual inspection of machine components. Expert Syst Appl 38(4):3260–3266
7. Sinha A, Yadav A, Gahlot A (2018) Extractive text summarization using neural networks. arXiv:1802.10137
8. Sanjabi N (2018) Abstractive text summarization with attention-based mechanism. Master's thesis, Universitat Polit_ecnica de Catalunya
9. Yogan JK, Goh OS, Halizah B, Ngo HC, Puspalata C (2016) A review on automatic text summarization approaches. J Comput Sci 12(4):178–190

# An Effective PUF Based Lightweight Authentication and Key Sharing Scheme for IoT Devices

**M. Prasanna Kumar, N. Nalini, and Prasad Naik Hamsavath**

## 1 Introduction

IoT is the future connected world of every physical object around us and these objects communicate with each other without human intervention to assist human beings in their various day to day activities. IoT has smart applications in various areas like medical, smart cities, industrial automation and so on. Trust is very important in the whole IoT ecosystem. The IoT networks become more complex with billions of connected devices across the globe, it is necessary to secure each of the elements ranging from IoT devices to connectivity to cloud is very important. The IoT with a loT of connected devices brings up a lot of issues and concerns about security. The IoT security brings some challenges that are relatively new and is important as it goes all the way to our devices. They are pulling out some data from that environment or controlling something. So, end to end security must be good. IoT is not just connecting devices, adding some level of intelligence. The attacks become more and more frequent when devices become connected. People are giving up some of the freedom for having the convenience of IoT.

Today a number of complex cryptographic solutions are available to provide secure communication in a network of the device. But using existing cryptography solutions is not feasible to adopt in constrained applications of IoT. Several technical obstacles, along with scarce resources like energy, computing power and storage is a challenge to address various IoT security requirements. For example, unauthorized access to IoT devices by using default credentials remains largely unsolved [1]. There exist a plethora of security solutions to enhance IoT security, many research and

M. Prasanna Kumar (✉)
Shree Siddhartha Institute of Technology, Tumakuru, India

N. Nalini · P. N. Hamsavath
Nitte Meenakshi Institute of Technology, Bengaluru, India
e-mail: nalini.n@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_22

operational problems yet to be solved, raising various concerns and thus hindering the trust in the IoT paradigm [1].

With respect to the device and hardware, the security and privacy threats include unauthorized device and network access with insecure external ports, lack of secure device configuration and software updates [2]. For example, in the smart city application scenario, accurate and reliable functioning of the IoT devices is the critical requirement for the decision-making systems to act on the data received. The authenticity of these devices must be assured to trust the decision-making process with very high confidence.

To summarize, the challenges of connecting personnel devices to the IoT network are

- Without proper access control on many IoT devices, adversaries can easily attack these devices and gain access to it using location information via the internet.
- IoT components usually communicate over wireless networks, an attacker could easily gain access to confidential information by eavesdropping.
- As most of the IoT devices operates on battery power with limited computing capability, IoT elements cannot support complex security schemes.
- Implementing centralized security has a threat of compromise attack and hence not feasible in IoT paradigm.

In such scenarios, designings secure mutual authentication protocol [3, 4] between two devices is a challenge that is affordable to resource constrained devices. One of the promising hardware-based solutions is to implement security mechanisms in devices with limited resources is the physically unclonable functions (PUFs). PUFs is a kind of digital fingerprinting technique that can be used to achieve things like authentication without using secret keys stored in a device. The PUFs hardware can generate unique challenge-response pairs (CRP) for each PUF enabled device. These salient features of PUFs are exploited in our authentication mechanism presented in this paper. Here an effective PUF based lightweight authentication and key sharing scheme is proposed. Unlike other schemes, our scheme uses a single CRP on the device and also the stored CRP dynamically changes during each communication session.

The rest of the paper is organized as follows. Some of the existing security mechanisms used in IoT and their limitations are discussed in Sect. 2. Section 3 describes the working of the proposed system. In Sect. 4 efficiency of our proposed system and uniqueness are discussed. Finally, the paper is concluded with the efficiency of the proposed work and the future enhancement of the last Sect. 5.

## 2 Related Works

Device authentication and controlling access to it [5–7] is one of the security challenges to be addressed in IoT systems. IoT systems use public internet as the backbone for operation, there are chances of misuse or damage to personal devices and personal information [8]. Two major challenges in this aspect are, the centralized approaches to manage authentication and access control becomes a bottleneck when the number of devices connected increase dramatically. As IoT devices have resource constraints, implementing traditional public key cryptographic algorithms is difficult. It demands a lightweight algorithm consuming less resources yet providing the same security level. Public key cryptography uses larger size keys which is not suitable for IoT devices as the devices have limited computational power and are difficult to process the keys. Reference [9] use dynamic session key to provide lightweight mutual authentication with block cypher algorithm. Elliptic curve cryptography produces shorter keys with a high level of security as compared to key of the same size produced by public key cryptography. Several ECC based authentication schemes are proposed for IoT [10, 11]. But if key is short it is easy to brute force. In [11] ECC is deployed with a pseudo random number generator for authentication of RFID tags and server.

To address the new privacy and security issues raised in the IoT ecosystem requires access control schemes to be built on the principles of high scalability, flexibility, lightweight and causality. Physically Unclonable Function (PUF) is a promising hardware technology that can be used to uniquely identify the devices connected in IoT. In PUFs, the outputs depend on the intrinsic characteristics of the PUF's physical elements and are therefore it is difficult to predict the output and quite impossible to clone the device. The PUFs are easy to construct and evaluate which suits the primitive security requirement of IoT devices. Using PUFs combined with other factors can provide a strong authentication scheme in constrained IoT applications. Reference [12] presents a two-factor authentication scheme for IoT devices addressing privacy and resource constraint. This scheme considers PUFs as one of the factors for authentication. A password or a shared secret key is used as the second factor. In [13] an authentication and key exchange protocol has been developed which uses the concept of PUFs, Keyed Hash and Identity based Encryption (IBE). The protocol eliminates the overhead on verifier to store challenge-response database of the PUF and the dependency of imposing security mechanism to keep it secret. However, the protocol needs optimization of the resources for encrypting frames and the side channel attacks have to be explored.

Identity based encryptions (IBE) [14] generates keys from a known unique public identifier. IBE use either a password or a digital certificate having the user's public key as a proof of identity instead of physical identity. However, some of these secret data need to be explicitly stored in the devices. Further, standard IBE uses Public Key Generator (PKG) which is used to generate private keys for the devices and transfer them through secured medium. The scalability of IoT applications makes key exchange for real life deployment of a difficult challenge. The scheme proposed in [9], uses Elliptic Curve Cryptosystem with PUFs for IOT framework, but as it stores helper data for each challenge in the device can lead to unacceptable memory overhead in the devices with scarce memory. Reference [15] proposes a PUFs based mutual authentication protocol but has a considerable hardware overhead; hence, it is not suitable for IoT paradigm. Moreover, many of the authentication schemes based on PUFs uses two elements a verifier device and a prover device. Verifier which grants authenticity to a prover device has a prestored subset of CRP data table or the instance of the PUF model embedded in the prover device.

## 3 Proposed System

The Physically Unclonable Functions are the promising hardware-based security technique that are more efficient to operate in a system with scarce resources. Hence PUFs proves to be a better choice to use in an IoT environment [9, 12, 13, 15, 16]. In this paper, PUFs unique features are exploited to implement an efficient authentication and session key generation mechanism. The system operates in two phases: *Initial CRP exchange phase and Authentication and session key sharing phase.* The first phase is executed in secured trusted medium. The devices and the server generate CRPs and exchange them with one another. The second phase process involves authentication and session key generation for devices in the IoT system.

### A. Initial CRP exchange phasee

PUFs based devices generate unique challenge-response pairs. Here the IoT devices are assumed to be PUF enabled. Here each device generates a challenge-response pair (CRP) and is exchanged with the challenge-response pair generated by the server. Figure 1 shows the process of initial CRP exchange between a device and the server. Here both will choose a random challenge and generate a response using PUFs. The device generates a challenge-response pair $\{C_d, R_d\}$ and sends to the server. Similarly, the server generates a challenge-response pair and exchanges with the device. Finally, the device stores the tuple $<C_d, C_s, R_s>$ and the server stores the tuple $<C_s, C_d, R_d>$ in their memory. This information stored is used for authentication and to establish a session key for communication in the next phase. The above phase is executed in a trusted medium.

### B. Authentication and session key sharing phase

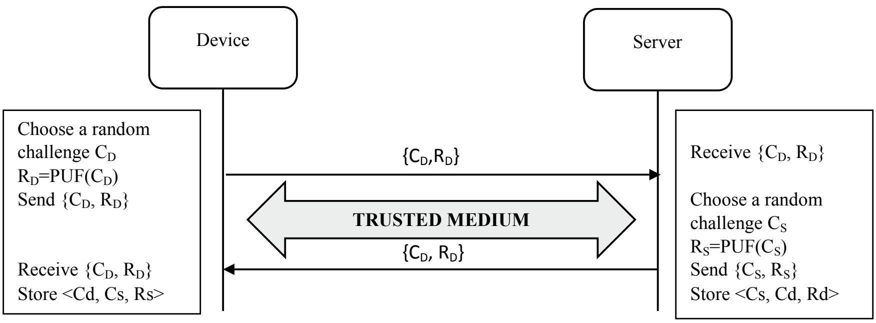



**Fig. 1** Initial CRP exchange phase

Figure 2 show the various steps involved in authenticating and sharing session key and the same is described below.

*Device request*: Initially, the device which wants to communicate with the server, sends a request with its credentials to authenticate and allow for communication. During the initial CRP exchange phase, the device has saved a tuple $<C_d, C_s, R_s>$ containing a CRP of sever and the challenge from the CRP it has shared with the server. From this data, the device first computes PUF output for the challenge $C_D$ on the go. The device then finds a key $K = h(R_d \| R_s \| C_s)$ and choose a random number Rn1. After these computations, the device sends a request message to the server with {K, Rn1}.

*Sever Response*: When the server gets the request message, the server first finds PUF output $R_S$ for the challenge $C_S$, i.e. $R_s = PUF(C_s)$. Using $R_S$ and $(C_S, R_D)$ from the saved data $<C_s, C_d, R_d>$, it computes the value $h(R_d \| R_s \| C_s)$ and verifies with

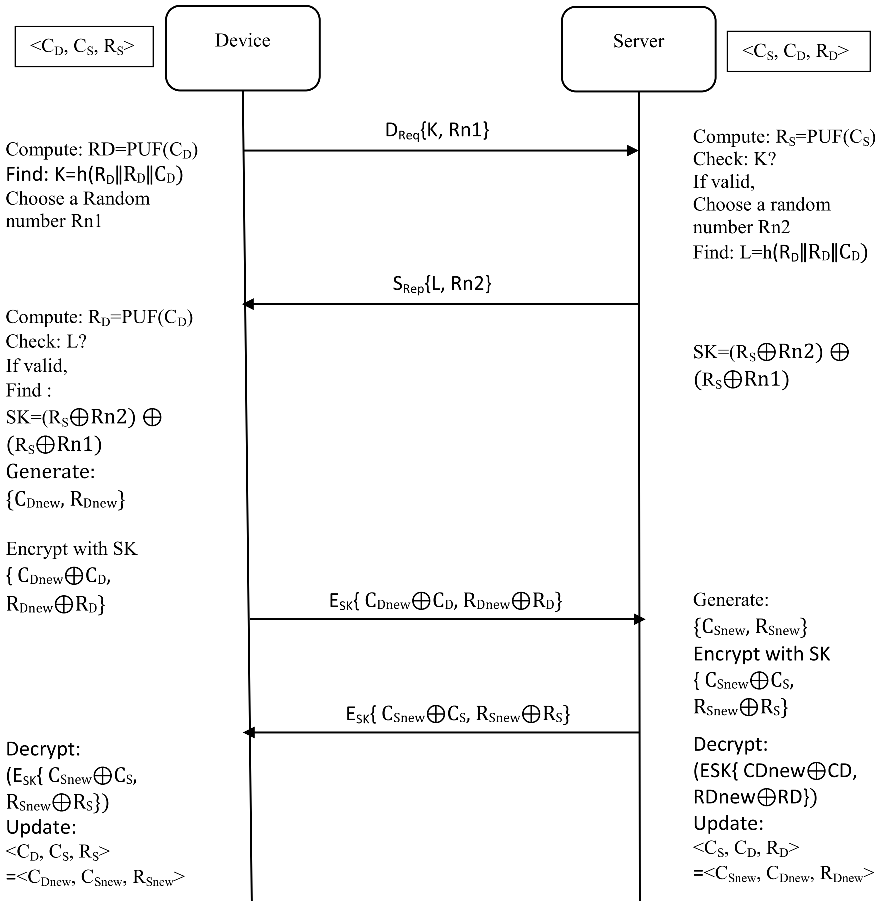



**Fig. 2** Authentication and session key sharing phase

the K. If it is valid, the device request is accepted, otherwise, the request is rejected. Now server choose a random number Rn2 and computes $L = h(R_d \| Rs \| C_d)$. Finally, the server sends {L, Rn2} in its response.

*Generating Session key and a new CRP(Device)*: On receiving the server response, the device validates the Value L with its available data. If it matches, the device will accept the server response. Then it computes and saves the session key $SK = (R_s \oplus Rn2) \oplus (R_S \oplus Rn1)$ for further communication. Now the device will generate a new CRP, $\{C_{Dnew}, R_{Dnew}\}$ and encrypts the message $\{C_{Dnew} \oplus C_D, R_{Dnew} \oplus R_D\}$ with SK and sends to the server for updating. The new CRP is xored with old CRP in the message to not to disclose the actual new CRP in any of the ways during transmission. So that even an adversary decrypt the message, actual CRP cannot be extracted.

*Generating Session key, new CRP and Updating* (*Server*): On receiving the above message sent, the sever first computes the session key and generates a new CRP for it. Then the server encrypts its new CRP using the session key, ESK $\{C_{Snew} \oplus C_S, R_{Snew} \oplus R_S\}$ and sends the message to the device for updating. Meanwhile, the server decrypts the message it has received and extracts new CRP of the device and updates with $<C_S, C_D, R_D> = <C_{Snew}, C_{Dnew}, R_{Dnew}>$.

*Device updating with new server CRP*: Now the device will decrypt the message sent in the previous step and extract the new CRP of the server. The same is updated with $<C_D, C_S, R_S> = <C_{Dnew}, C_{Snew}, R_{Snew}>$.

## 4 Discussion

The PUF based security mechanism used for authentication in the presented system has proven to be more efficient. Since hardware is a part of the security mechanism here, which speeds up the execution of the protocol. The hardware implementation and the PUF unique challenge-response generation for each device overcome the drawback of a single key stored in the device. Our proposed system performance is evaluated with two recent PUF based authentication schemes Chatterjee et al. [13] and Gope et al. [12]. Summary of the comparison of our scheme with the other two is shown in Table 1. It clearly shows that our scheme is resilient to any password guessing attack, no multifactor authentication complexity. Our scheme won't use any verifier to maintain CRP tables. Hence there is no chance of compromise of the verifier. The proposed system is simple and efficient against many common attacks and stands unique in not maintaining CRP tables and also the CRP used during the authentication process is dynamical changes for each communication session.

**Table 1** Performance comparison with existing PUF based protocols

| Parameters | Gope and Sikdar [12] | Chatterjee et al. [13] | Proposed system |
|---|---|---|---|
| Resilience to man in the middle attack | Yes | Yes | Yes |
| Resilience to physical access | Yes | Yes | Yes |
| Resilience to password guessing attack | No | Yes | Yes |
| No multi factor authentication complexity | No | No | Yes |
| Resilience to verifier compromise | Yes | No | Yes |
| Algorithm complexity | Medium | High | Low |
| No maintenance of CRP table | No | No | Yes |

## 5 Conclusion

PUF is a novel hardware technique for many security applications. PUF based security solutions for IoT devices of limited resources is more efficient. In this paper, an effective PUF based authentication mechanism is implemented for IoT applications. The proposed mechanism is secure against many of the common security attacks in IoT environment. Our system stores minimum data in the devices and also no secret data is directly or in an encrypted manner is shared in the channel during the authentication and session key generation process. In future advancements, the security mechanism presented in the paper can be improved with multi factor authentication and other efficient cryptographic techniques like elliptic curve cryptography.

## References


1. Neshenko N, Bou-Harb E, Crichigno J, Kaddoum G, Ghani N (2019) Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations. IEEE Commun Surv Tutorials 21(3):2702–2733
2. Li J, Yan Q, Chang V (2018) Internet of things: security and privacy in a connected world. Future Gener Comput Syst 78:931–932. https://doi.org/10.1016/j.future.2017.09.017
3. Li N, Liu D, Nepal S (2017) Lightweight mutual authentication for IoT and its applications. IEEE Trans Sustain Comput 2(4):359–370. https://doi.org/10.1109/TSUSC.2017.2716953
4. Qiu Y, Ma M (2016) A mutual authentication and key establishment scheme for M2M communication in 6LoWPAN networks. IEEE Trans Industr Inf 12(6):2074–2085. https://doi.org/10.1109/TII.2016.2604681
5. Ferdowsi A, Saad W (2019) Deep learning for signal authentication and security in massive internet-of-things systems. IEEE Trans Commun 67(2):1371–1387. https://doi.org/10.1109/TCOMM.2018.2878025
6. Lin S, Wen C, Sethares WA (2018) Two-tier device-based authentication protocol against PUEA attacks for IoT applications. IEEE Trans Sig Inf Process Over Networks 4(1):33–47. https://doi.org/10.1109/TSIPN.2017.2723761

7. Daddala B, Wang H, Javaid AY (2017)Design and implementation of a customized encryption algorithm for authentication and secure communication between devices. In: IEEE national aerospace and electronics conference (NAECON), Dayton, OH, pp 258–262. https://doi.org/10.1109/NAECON.2017.8268781
8. Yu Y, Li Y, Tian J, Liu J (2018) Blockchain-based solutions to security and privacy issues in the internet of things. IEEE Wirel Commun 25(6):12–18. https://doi.org/10.1109/MWC.2017.1800116
9. Wallrabenstein JR (2016) Practical and secure IoT device authentication using physical unclonable functions. In: IEEE 4th international conference on future internet of things and cloud (FiCloud), pp 99–106
10. He D, Zeadally S (2015) An analysis of RFID authentication schemes for internet of things in healthcare environment using elliptic curve cryptography. IEEE Internet Things J 2(1):72–83. https://doi.org/10.1109/JIOT.2014.2360121
11. Tewari A, Gupta BB (2018)A mutual authentication protocol for IoT devices using elliptic curve cryptography. In: 8th International conference on cloud computing, data science and engineering (confluence), Noida, pp 716–720. https://doi.org/10.1109/CONFLUENCE.2018.8442962
12. Gope P, Sikdar B (2019) Lightweight and privacy-preserving two-factor authentication scheme for IoT devices. IEEE Internet Things J 6(1):580–589. https://doi.org/10.1109/JIOT.2018.2846299
13. Chatterjee U et al (2019) Building PUF based authentication and key exchange protocol for IoT without explicit CRPs in verifier database. IEEE Trans. Dependable Secure Comput 16(3):424–437. https://doi.org/10.1109/TDSC.2018.2832201
14. Chen W (2012)An IBE-based security scheme on Internet of Things. In: IEEE 2nd international conference on cloud computing and intelligence systems, Hangzhou, pp 1046–1049. https://doi.org/10.1109/CCIS.2012.6664541
15. Che W, Martin M, Pocklassery G, Kajuluri VK, Saqib F, Plusquellic J (2016) A privacy-preserving, mutual puf-based authentication protocol. Cryptography 1(1)
16. Mukhopadhyay D (2016) PUFs as promising tools for security in internet of things. IEEE Des Test 33(3):103–115. https://doi.org/10.1109/MDAT.2016.2544845

# IoT-CBSE: A Search Engine for Semantic Internet of Things

**R. Raghu Nandan, N. Nalini, and Prasad Naik Hamsavath**

## 1 Introduction

The Internet of Things (IoT) is an ecosystem of interconnection of physical objects, computing devices, electronic gadgets, sensors and actuators, cars and trucks, machinery in industries and utility components, and any day to day objects with Internet connection with the capability of data analytics and have the power of gathering and exchange of the gathered information.

As the IoT ecosystem is elaborating, the devices should possess the ability to discover their target devices with predefined device characteristics, and autonomously sync with the neighboring devices to accomplish useful tasks. A smart device discovery should prompt the IoT devices to effectively establish fresh connections with neighboring devices based on their necessity.

A large number of computing capabilities, network services, and data may be combined with the assistance of the smart device discovery mechanism. For example, Smart agriculture is used to represent the application of IoT for the optimal yield in agriculture. By deploying the sensors in the farmland, the farmers can gather information about the crop environment and can make intelligent decisions so as to increase the yield of the crop and also the livestock maintenance. By the way of precision farming agriculturists can supervise the state of growth stage of the crop, and will be able to make suitable decisions on the amount of spraying pesticides so as to obtain the optimal yield. In this scenario, an intelligent device discovery could

R. Raghu Nandan (✉) · N. Nalini · P. N. Hamsavath (✉)
Navkis College of Engineering, Hassan, India
e-mail: naikphd@gmail.com

N. Nalini
e-mail: nalini.n@nmit.ac.in

P. N. Hamsavath
Nitte Meenakshi Institute of Technology, Bengaluru, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_23

help the agriculturist to reach the correct computing facilities, network services and data analytics points to extract appropriate decisions about the next activity in the cropland. In addition, the IoT devices can form the social connections as human beings do, and they can combine intelligence of smart devices, and data, without the intervention of human beings. In this case, an intelligent device discovery should have a forwarding strategy that guides the discovery request to quickly arrive at the desired target device with minimum detours. In addition, since most of the IoT devices are power constrained, the device discovery should avoid involving too many intermediate devices in the process of data exchange.

## 2 Literature Review

Marchal et al. [1] developed a self learning search engine AUDI that operates on a distributed system for learning and discovering the different categories of IoT devices. A novel device type discovery technique based upon passive fingerprinting of episodic communication traffic of IoT devices. It doesn't require any earlier knowledge of device types or the labeled training data and is efficient in discovering IoT device types in every operation mode of a device. AUDI creates abstract device type labels that can be used autonomously operating systems without human intervention.

Asokan et al. [2] DÏoT: An autonomous System for identifying Compromised IoT Devices. The proposed solution depends on novel automated techniques for device category recognition and device type-specific anomaly detection. It identifies compromised IoT devices quickly (≈2 s) and accurately (success rate of 94%), without raising any false alarm in a period of one week of evaluation in a real world deployment setup.

Tippenhauer et al. [3] proposes a method for classifying IoT devices along with their manufacturer name and year, with 99.281% accuracy that assures proposed method is suitable for real enterprises. The novel approach might be used to automatically and accurately identify connect IoT devices in an organization's computer network, and thus decrease violations of operational policies. The advanced scenarios may possibly include different network rules and various data capturing points.

Kamilaris et al. [4] developed an approach for discovering WoT devices and services and semantically annotate data present on the web of the WoT devices, by making use of excellent web crawlers acting as agents for a discovery search engine on SWOT. WOT Semantic search Engine (WOTS2E) is a real-time universally accepted search engine for Semantic WoT, capable of discovering Wot-enabled devices and their services.

Miettinen et al. [5] proposes a framework that performs device identification on the basis of profiling of the device type specific communication behavior of individual devices. The proposed method gives security for user's network by holding network separation, in which communications of vulnerable devices are completely avoided, thereby effectively controlling security risks related to the devices. The framework applies software-defined networking to perform network isolation and traffic filtering. The authors also put forth that the enforcement framework can be efficiently implemented with moderate impact on traffic latency.

Kotak et al. [6] presented a method that makes use of deep learning to discover both known and unauthorized IoT devices in the network traffic, identifying 10 different IoT devices and the traffic of smart phones and computers, with over 99% accuracy, and achieving over 99% overall average accuracy to detect unauthorized IoT devices connected to the network. The proposed methodology is generic as it focuses on the network traffic payload of the different IoT devices as against the header of the packets; thus the method is applicable to any IoT device, irrespective of the network protocol used for the exchange of information.

Akhileshwari et al. [7] proposes a framework that mandates the devices and their services to be registered ahead of their communication, the algorithm fetches the appropriate device and service for the search request entered by the user by the method of semantic search. The framework adopts a semantic searching strategy to search for appropriate data and also match and bind the exact device with a well-suited service.

Lo et al. [8] presents a specific algorithm by name You Only Look Once (YOLO) that is an object detection algorithm that uses the unique time-frequency visualization of LoRa signals. YOLO is a simple method of IoT device discovery. With intensive analysis of the problem domain and extensive image augmentation and simulation, the obtained discovery system is capable of learning general representation of these signal objects. The drawback is that humans require different resolution images to differentiate all LoRavariants.

Sunthonlap et al. [9] the proposed Social-Aware and Distributed (SAND) method employs the Depth First Search (DFS) technique and sends the discovery request to neighboring devices in a ranked order defined by social network-aware device ranking strategy. Usually, the ranking criteria take into account the device's degree, diversity and clustering coefficient, which could specifically route the discovery request to reach a wide area. Furthermore, the local betweenness is measured in the ranking criteria, which prioritizes the devices that stand at the critical intersections of multiple shortest paths of the network.

## 3 Problem Identification

The major hurdle in the Internet of Things (IoT) is to identify the exact device, which leads to having a successful communication. In the present growing Internet of Things ecosystem, it is desirable to have a fast and accurate device discovery search engine to have a trusted and appropriate communication and which is the need of the hour. In the present condition, the search engines cannot be trained on the preprocessed data as the growth of the IoT is exponential and various new categories of devices are getting attached to the network every hour. Hence there is a need to develop a search engine that learns autonomously from the description of the user and locates the appropriate URI of the devices connected to the internet, that intern helps in hassle free communication by the intended user. Therefore, our aim is to develop a semantic search engine that autonomously identifies the devices in the IoT ecosystem and returns the appropriate device identification.

## 4 Contribution and Organization of the Research

The proposed system includes the following contributions:

1. The proposed search engine is developed for device discovery in the Internet of Things Ecosystem.
2. The proposed framework is optimized for faster access and accurate device discovery.
3. The proposed framework is self learning or autonomous in nature for device discovery.

This paper is organized as Sect. 1 gives the introduction, Sect. 2 Literature review, Sect. 4 contribution, Sect. 5 the proposed search engine for device discovery and Sect. 6 gives the conclusion.

## 5 Search Engine for Device Discovery

Search engines play a major role in the evolution of the Internet of Things (IoT) as these devices are connected to the network as well as to themselves in the ecosystem. The devices keep on gathering information about the surroundings and updating database continuously. At the database, one can extract the data generated by the device as well as the data on the device can also be extracted. The data on the device is said to be the metadata on the device which gives the description of the devices.

With the increase in a number of connected devices abundant amounts of generated data is available on the connected environment, discovery of the devices and metadata on devices gaining much importance and also poses a challenge to extract the metadata. In the past, consumers of sensor data normally also fulfilled the role of data generation with a strong tie between application and sensor network. This tie has started to break, as the data generators now are unaware of where and how their data is being used, and data consumers are unaware of where and how to find the relevant data about the data generators.

The Sensor Model Language (SensorML) does the pulling out of the metadata on the coordinate based geometric characteristics of sensors and sensor systems. By this data on data, it becomes possible to identify the geo-spatial coordinates in which a device works; it requires an extra step to verify whether the particular coordinates of each sensor come within the user's target location.

The data generated by these devices can be formed in the form of linked data. Linked data refers to data published on the web in such a way that it is machine

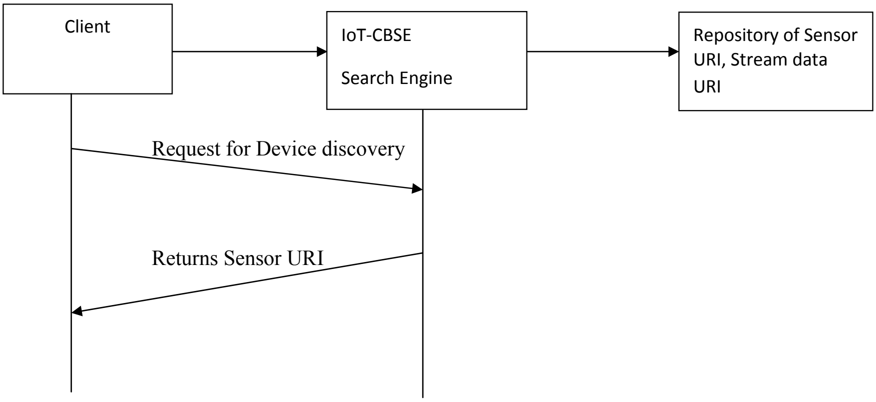



**Fig. 1** IoT-CBSE search engine

readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked to from the external data sets.

The proposed system uses sensor URI for device discovery. The data returned from a sensor URI should be the metadata about the sensor and from the stream data, URI should be the observation of the sensor.

Figure 1 depicts the operation of the Internet of Things Clue Based Search Engine (IoT-CBSE). In the figure, the client is one who wants to establish communication with the devices in the IoT environment can request the search engine to identify the appropriate device by describing the nature of the device or few words about the device operation. The search engine searches for the repository of device descriptions through URIs, these URIs are classified with the specific class of devices based on the metadata of the devices, this metadata gives the description about the devices which helps in identifying the devices.

Whenever the client requests the device discovery, the search engine self learns about the description given by the client and searches the repository of the URIs, and find the appropriate URI that matches the description given by the client, after finding the appropriate URI that URI is sent in the form of RESTful API, which is the Representation State Transfer, then the client can extract the communication address of the intended devices and can communicate with that device and fulfill the needs of search of the client. Representational State Transfer is the architectural style for building loosely coupled web applications over the internet. The proposed system is one of the best device discovery systems based on the clue present in the description of the device.

## 6 Conclusion

In the fast and vast growing IoT ecosystem, there is a need for search engines that can operate at the user interaction with no much technical knowledge of the Internet of Things. The client will request the search engine by describing the nature of the sensor or the operation of the sensor. The search engine should accurately grab the semantic meaning of the description and should be able to accurately discover the device of interest of the client and redirect the client to the appropriate device to have meaningful communication. Hence our proposed system fulfills the needs of the clients who are in search of the appropriate device but not clearly able to technically describe the device, IoT-CBSE will help such clients to locate the appropriate device by using a wise device discovery strategy by analyzing the natural description by the client and providing accurate URI for their communication.

## 7 Future Enhancements

Since our work is in the initial stage, we have presented a search engine based on semantic search. In the future, we will enhance the search engine with advanced Machine Learning and Natural Language Processing that enables the search engine useful for advanced users for the purpose of research, since we are intended to include ML and NLP.

## References

1. Marchal S, Miettinen M, Nguyen TD, Sadeghi A-R, Asokan N (2019) AUDI: towards autonomous IoT device-type identification using periodic communication. IEEE J Sel Areas Commun 37(6):1402–1412. https://doi.org/10.1109/JSAC.2019.2904364
2. Nguyen TD, Marchal S, Miettinen M, Asokan N, Sadeghi A (2018) DÏoT: a self-learning system for detecting compromised IoT devices. CoRR, vol abs/1804.07474
3. Meidan Y, Bohadana M, Shabtai A, Guarnizo JD, Ochoa M, Tippenhauer NO, Elovici1 Y (2017) ProfilIoT: a machine learning approach for IoT device identification based on network traffic analysis. SAC'17, Marrakech, Morocco. https://doi.org/10.1145/3019612.3019878
4. Kamilaris A, Yumusak S, Ali MI (2016) WOTS2E: a search engine for a semantic web of things. In: IEEE 3rd world forum on internet of things (WF-IoT), Reston, VA, pp 436–441. https://doi.org/10.1109/WF-IoT.2016.7845448
5. Miettinen M, Sadeghi A-R, Marchal S, Asokan N, Hafeez I, Tarkoma S (2016) IOT SENTINEL: automated device-type identification for security enforcement in IoT. arXiv:1611.04880v2 [cs.CR]
6. Kotak J, Elovici Y IoT device identification using deep learning. arXiv:2002.11686v1 [cs.CR]
7. Akhileshwari KG, Salian S (2019) IoT device and service discovery framework. Int J Recent Technol Eng (IJRTE) 8(2S6). ISSN: 2277–3878
8. Lo E, Kohl J (2020) Internet of Things (IoT) discovery using deep neural networks. In: IEEE winter conference on applications of computer vision (WACV), Snowmass Village, CO, USA, 2020, pp 795–803. https://doi.org/10.1109/WACV45572.2020.9093371

9. Sunthonlap J, Nguyen P, Ye Z (2018) Intelligent device discovery in the internet of things—enabling the robot society. arXiv:1712.08296v2 [cs.AI]

# Flood Monitoring and Alerting System for Low Lying Urban Areas

**S. Pradeep Reddy, T. R. Vinay, K. Manasa, D. V. Mahalakshmi, S. Sandeep, and V. Muthuraju**

## 1 Introduction

The temporary overflow of water onto land which is usually dry is known as flooding. It occurs in almost all types from ephemeral streams to dry channels and from humid zones to arid climates and in rivers, streams, and channels. The floods in urban areas caused a great amount of destruction to roads, bridges, electric poles, and inconvenience such as paralyzing day to day activities, street water logging which in turn causes a heavy traffic jam, leading to severe economic damage.

Though cyclones can be tracked using satellite images, it is necessary to have real-time data such as water level to make a decision that is to prevent flooding, temperature, flow, pressure, humidity, and precipitation level. Huge damage to properties due to floods involves a huge amount of loss to the government and individuals in many countries in the world, hence an efficient operating system for flood response between different agencies is necessary to monitor activities. Warning that is being given before flood event has direct effect on reducing the damage caused by floods, hence flood monitoring and an alerting system will reduce the amount of money spent on the problem and loss of life can be reduced and carried out way quicker if a safe region is known.

The field of computer science and applied electronics where the system of devices collects real-time data and transfers it through a wireless sensor network is termed as Internet of things, specifically we use water sensors to measure the rate of water flow or to measure the water level and we use GSM module to send the warning messages to the people and various rescue teams about the rising water level. For flood prediction, techniques such as Machine Learning are used for prediction and visualizing the data available and finding out the area in a region where the flooding might happen and an area where the people can temporarily migrate to.

---

S. Pradeep Reddy (✉) · T. R. Vinay · K. Manasa · D. V. Mahalakshmi · S. Sandeep · V. Muthuraju
Nitte Meenakshi Institute of Technology, Bengaluru, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_24

In an emerging world with advanced technologies, the proposed system involves three stages wherein an initial stage involves the measure of water level through ultrasonic sensor [1] and prediction of flood status value. Secondly, to alert the public in remote areas and authorities via SMS through GSM modem. Thirdly, retrieving real-time data and depicting the flood-prone zones via an android application.

## 2 Related Work

As per the study, In the already existing system which uses no-Wan model data parameters are collected physically, and then the prediction is done [2]. And in another paper, they just had the idea to do prediction which was to use Arduino [3] and sensor, but was in the idea stage and it was not developed for deployment. And in a few other papers had the model done with sensors that were not tested. And the use of Machine Learning Algorithms to predict but it was not accurate. The existing approach makes use of the WSN model along with statistical data for flood prediction and forecasting Machine Learning is being used. A recent study on flood prediction revealed the extensive use of the Wireless Sensor Network [4] and advanced artificial network. Earlier non-WSN based approaches like hydrologic models were used to predict floods. A physical detection system in hydrologic models to predict floods and it was done based on obtained parameters [5]. In 1970s Bellandur lake, one of the oldest and largest lakes of Bangalore which has been located near Bellandur which serves 18 villages in the surrounding decades ago.

People in those villages used to harvest paddy and grow various vegetables for example tomatoes and cauliflower abundantly and people around the lake used to carry out fishing. More than 400 families used to live in the neighborhood. The water from that lake was so pure it was used as drinking water as well as for the irrigation was carried out. The people and nearby villages to that lake used to commemorate the annual boat festival which was named Theppotsava.

According to the Lake Development Authority (LDA) decades ago the lake was spread across more than 890.09 acres of land. Day–by-day lake size got diminished because of urbanization and now the lake covers 726 acres only (Fig. 1).

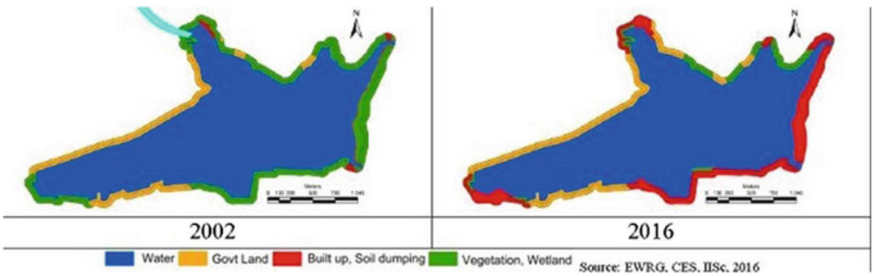



**Fig. 1** Bellandur Lake


**Fig. 2** Flood disaster

Bellandur Lake is one of the most polluted lakes in Bengaluru and it's part of the Varthur lake series. This lake has more inlets that serve polluted water from the industries. Because of the untreated sewage water entering this lake, it is now nothing less than a sewage tank. According to records, clean and unpolluted rainwater from the valleys of Challaghatta and Koramangala through various stormwater drains should flow to Bellandur lake. However, sewage inflow to this lake is due to two reasons. (1) Overflow of untreated sewage water due to Improper sewage system from the factories joins stormwater drains and merges with rainwater from the valleys of Challaghatta and Koramangala which in turn flows to Bellandur Lake. (2) Lack of sewage treatment plants to treat wastes coming out from the small industries which also gets mixed with the stormwater drains.

According to a research paper written by Han et al. [5] presented an idea called the Hydrology monitoring system using a technology called ZigBee along with the star topology. But in their work, there were no explicit flood prediction schemes. Ancona et al. [6] described an approach that included IoT with sensors, viz river, and rain gauges, which were quite efficient energy-wise, and technology like Wi-Fi was used for communicating (Fig. 2).

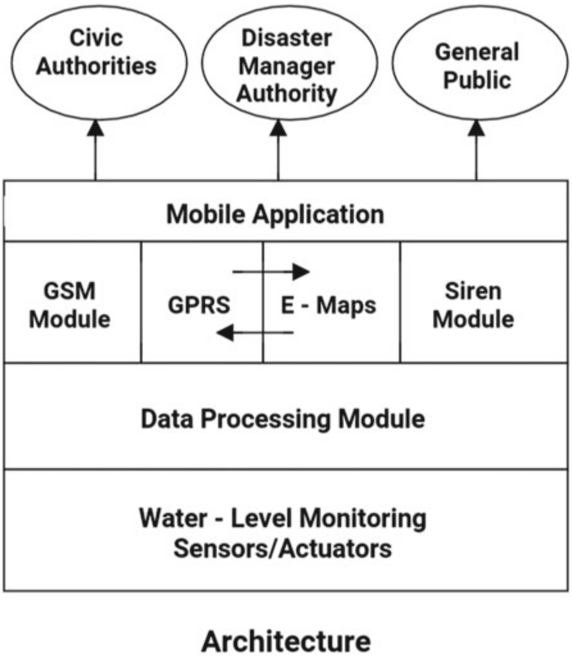
As per the study, In the already existing system which uses no-Wan model data parameters are collected physically, and then the prediction is done. And in another paper, they just had the idea to do prediction which was to use Arduino and sensor, but it was in the idea stage and it was not developed for deployment. And in a few other papers had the model done with sensors that were not tested. And the use of Machine Learning Algorithms to predict but was not Accurate. The existing approach makes use of the WSN model along with statistical data for flood prediction and forecasting Machine Learning is being used. A recent study on flood prediction revealed the extensive use of the Wireless Sensor Network and advanced artificial network.

## 3 Architectural Framework

The monitoring device is composed of an ultrasonic sensor which is used to measure water level in water bodies, an Arduino micro-controller to process the signal from the sensor, GSM module to transfer data or information from an Arduino to the computer server, and power in the form of battery and Regulator. As soon as the sensor is triggered, an output signal will be relayed on an Arduino micro-controller which acts as the switch and triggers the connected GSM module to send an alert message and water level status to the other GSM module on the connected computer server. The developed system keeps uploading all the readings and flood status into the cloud and has been automatically displayed on the android application developed. On uploading, if the flood status indicates the danger level crosses the threshold, the program will then automatically broadcast the alert message to the people. Furthermore, the concerned authorities or local officers can get the status and values of water level just by logging into the android application which is being developed. The developed system also includes a feature of alerting the public through the sound buzzer system (Fig. 3).

On many coastal regions and river valleys, ultrasonic sensors have been deployed to provide data about tropical storm surge. To monitor flooding, they have also been distributed on similar platforms on different rivers. The recent alerting system is being implemented on an urban street to monitor floods. We opted ultrasonic sensor for the project since ultrasonic sensing has been continuously developed for water

**Fig. 3** Architectural diagram



level measurement. The main duty of GSM in this system is to send SMS and it is the mode of communication between the maintenance authority and concerned authority. GSM module is interfaced with processed data from Arduino, so if it reaches the threshold value then Arduino triggers the GSM module to send SMS. As we know SMS is the cheapest way to communicate and it is also available in all types of mobiles, it is more reliable to reach a greater number of public. So, information dissemination can be achieved easily. Technically, the study tells us that apart from being the cheapest way available to transmit information to remote areas advantage of bandwidth usage also comes into the picture that is it does not require high data bandwidth and also SMS does not require cellular data to transmit, slight strength in the signal can do the work.

# 4 Methodology

## *4.1 Data Flow Diagram*

Water level data is being collected using an ultrasonic sensor and once collected the data is sent to cloud and the values are checked if it lies within the threshold range. If the value reaches the threshold there are different levels of threshold namely (Fig. 4):

- Low and Moderate level: the authorities are informed about it so that they can take necessary actions and precautions like increasing the outflow which would reduce the risk.
- Critical and High level: the warning message is sent to all the residents.

Safe areas are the ones whose altitude is higher than 950 m which are plotted on the map and made visible to users through mobile interface.

## *4.2 Algorithm*

**Algorithm 1: Prediction System**

1. Start
2. Read the water level values.
3. Generate fuzzy membership values.
4. Water level membership values.
5. Rainfall membership values.
6. Apply the activation function for these values.
7. Rules are formed using the rules evaluation.
8. Rules are aggregated.
9. Calculate the defuzzied value using centroid method that is flood status.

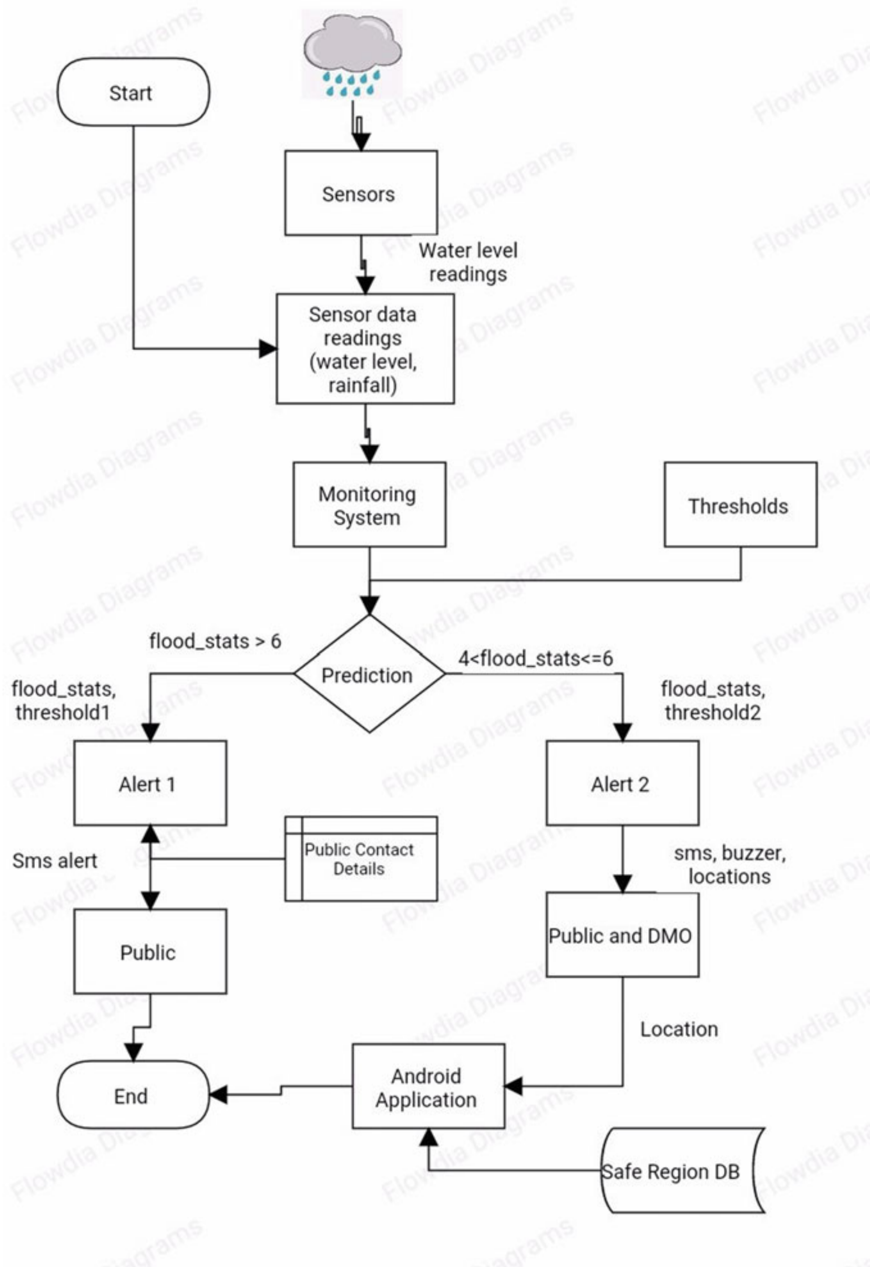



**Fig. 4** Data flow diagram

10. Return the flood status to Arduino.
11. Stop

**Algorithm 2: Alert System**

1. Start
2. Connect the components as required.
3. Initialize the trigger, echo pin of ultrasonic, and pin for buzzer.
4. Start measuring the readings from ultrasonic.
5. Send the readings to prediction system.
6. Receive the flood status from prediction system.
7. If flood status level is less than safe range does nothing.
8. If flood status level is in critical range.
9. Send alerts to the public through SMS and buzzer.
10. Repeat the procedure from step 3.
11. Stop

## 5 Results

The developed system involves the following test cases:

Table 1 depicts the recordings of measures of water level based on which they are categorized as normal, moderate, critical, and emergency. According to the alert level, SMS notifications are sent to the public and authorities such as disaster management systems, respectively.

Table 2 depicts the categories of alert level based on which flood warning status is being classified as safe, prepare for evacuation, and evacuate and SMS notification is delivered accordingly to public and disaster management authority.

**Table 1** Alert level category based on water level

| Water level (inch.) | Alert level | SMS notification delivery |
|---|---|---|
| 5 and below | Normal | 10-min interval |
| 6–10 | Moderate | 5-min interval |
| 11–15 | Critical | 1-min interval |
| 16 and above | Emergency | 1-min interval |

**Table 2** Flood warning status based on alert category

| Alert level | Flood warning status | SMS notification delivery |
|---|---|---|
| Normal | Safe | To disaster monitoring organization |
| Moderate | Prepare for evacuation | To disaster monitoring organization |
| Critical | Evacuate | To public and disaster monitoring organization |
| Emergency | Evacuate | To public |

Figures 5 and 6 depicts the functioning of application to monitor floods and alert people according to water level and flood status value. Dashboard displays the parameters like temperature, humidity, pressure, and water level measuring through ultrasonic sensor and calculated flood status value, in the map the flood pruned zones are shown using red marker and blue marker indicates the safe place for temporary migration whereas the blue circle indicates the radius about 2 km which can be varied up to 4 km. Current location of user is directly taken else can be entered manually accordingly with the user's convenience.

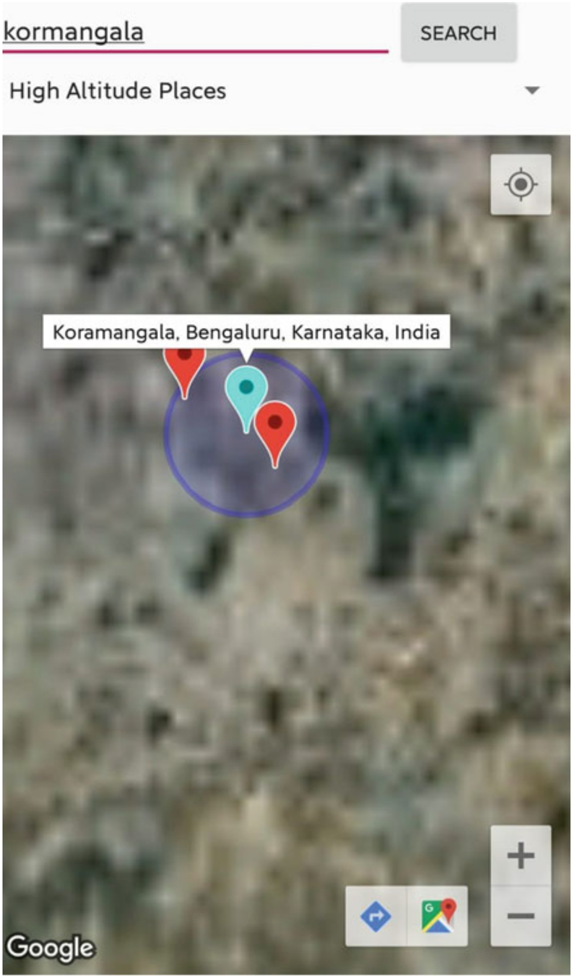
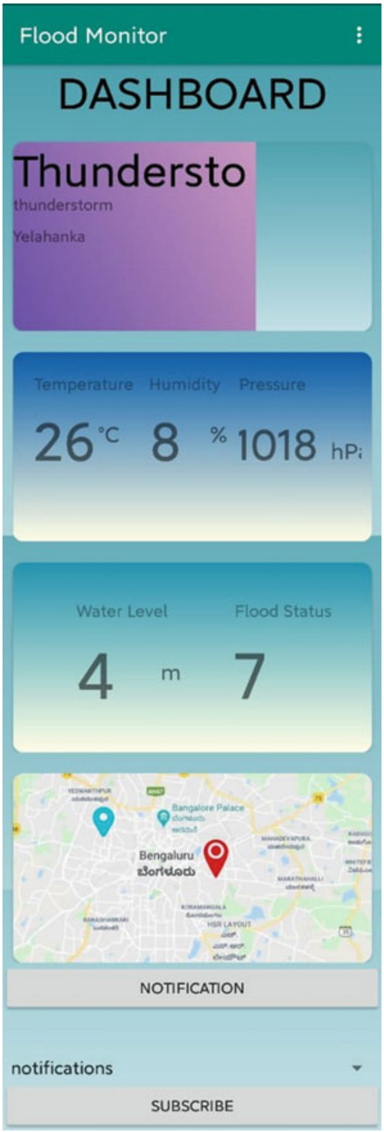



**Fig. 5** Map region

**Fig. 6** Dashboard




## 6 Conclusion

Due to the recent devastating flood faced by India in Belagavi, the need for efficient flood monitoring and alerting system arose. Therefore, effective ways to minimize the risk of flooding are issuing warnings about the flood and flood forecasting to the public and authorities.

The developed system work enlightening the reduced effect of flood risk through an alert system and also ensures increased accessibility and enhances efficiency in response to catastrophic floods, flash floods, enhanced effectiveness, and assessment of the emergency circumstances, which in turn can contribute to government authorities who help the society about providing information about floods. The warning message about flood prediction can be sent via system and the location of the safe area to the public and government authorities so that the necessary precautionary measures can be taken like temporary migration.

## References

1. Ultrasonic Sensors Introduction to sensors. https://www.keyence.com/ss/products/sensor/sensorbasics/ultra/index.jsp. Accessed 02 Nov 2019
2. Wu M, Lu T-J, Ling F-Y, Sun J, Du H-Y (2010) Research on the architecture of Internet of Things. In: 3rd International conference on advanced computer theory and engineering (ICACTE)
3. Mosavi A, Ozturk P, Chau K-W (2018) Flood prediction using machine learning models: literature review. Water 10:1136
4. Seal V, Raha A, Maity S, Mitra S, Mukherjee A, Kanti Naskar M (2012) A simple flood forecasting scheme using wireless sensor networks. Int J Ad hoc Sens Ubiq Comput (IJASUC) 3(1)
5. Beven KJ, Kirkby MJ, Schofield N, Tagg AF (1984) Testing a physically-based flood forecasting model (TOP-MODEL) for three UK catchments. J Hydrol 69(1–4):119–143
6. Ancona M, Dellacasa A, Delzanno G, Dibris A (2015) An "Internet of Things" vision of the flood monitoring problem. In: AMBIENT 2015: The fifth international conference on ambient computing, applications, services and technologies, pp 26–29. ISBN:978-1-61208-421-3

# Automatic Gate Control System

**V. Nishchay, P. Sujith Bhatt, S. Sreehari, M. N. Thippeswamy, and Dipak Kumar Bhagat**

## 1 Introduction

The ANPR is used to detect the region of a vehicle which consists of the number plate and could also be used to store snaps got from the camera and also the information in a way that a string which is got from the number plate and with a small tweak in the ordering it can also store the snap of the driver. Most systems use infrared cameras to capture snaps at any given point in time [1]. The exact implementation of ANPR technology differs at different places as a number of factors such as font color, size, and placement of the number plate have to be taken into consideration.

Radio-frequency identification (RFID) uses electromagnetic fields to automatically track and identify tags that are attached to particular objects. The tags contain particulars that are stored electronically. Two types of RFID tags, passive tags operate by collecting signals from a close RFID reader's interrogating radio waves and active tags use a localized power source to function and may operate hundreds of meters from the RFID reader. The RFID tags do not need to be within the line of sight of the reader, unlike bar code, so it may be embedded in the tracked object. RFID is one of the several methods of Automatic Identification and Data Capture (AIDC) [2].

Radio-frequency identification tags also known as RFID tags are utilized in multiple sectors. For instance, an RFID tag could be affixed to a car which in turn can be used to trail a progression through the assembly line, during production.

In [3], a detailed RFID based gate entrance set up is proposed for the university. The ID cards of the students themselves are used as there will be an existing database containing the information of the student. A valid card will allow the student to enter

V. Nishchay · P. Sujith Bhatt (✉) · S. Sreehari · M. N. Thippeswamy
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India

D. K. Bhagat
Subex Limited, Bengaluru, Karnataka, India

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_25

the campus and in case of an invalid card due to some reason, the pupil will not be authorized to enter the grounds. Passive RFID was used with MiF are RFID reader. RFID reader box was set up for students to tap the student ID to enter and exit the campus. It was also equipped with led, green led to show valid ID and a red led for invalid ID. Separate lanes were used for motorcycles and cars. The implemented system had a wait time of around 3–8 s for motorcycles and 5–10 s for cars.

A system for automatic vehicle detection and driver identification to go in or pull out a parking area is proposed in [4]. The authors develop a complete system which consists of three essential modules. Vehicle detection module detects the presence of vehicles by using a vehicle detection algorithm. Vehicles were successfully detected on images which contained more than one vehicle at different positions. The face detection module detects the driver's face after the vehicle has been detected. The authors achieved an average accuracy of 97.03%. Face recognition is a segment that recognizes the face of the driver. For this only a single gallery face image was present which was used for training. For the test images, four dissimilar drivers pose facial images ranging from frontal to +45 were used. They successfully achieved a complete system execution time of 0.5193 s.

QR code refers to a machine-readable code which is a combination of white and black squares [5, 6]. The use of QR code for automatic gate control system is detailed in [7]. The essential lead of using QR code is that it can be read by the scanner even if it is partially damaged and is easy to generate. The system is implemented by using Arduino Uno micro controller. The authors also developed an application for monitoring and controlling the system. The application was developed in VB.NET language. The system had an overall accuracy rate of 99%. The developed system is cost-efficient but only provides medium security. A parking gate control system controlled using a mobile application is detailed in [8]. Automatic barrier system was developed using Arduino Uno micro controller. The launch and termination of the barrier are controlled utilizing a mobile application preferably made on android. A Bluetooth module was used to achieve this wireless control. The android application has a button to open and close the barrier. On pressing the button, a command is sent which is used to revolve the mechanism and close or open the barrier. An IR sensor was also used to prevent the barrier from closing in case there is a vehicle still passing. The automatic check-post and tolling system in [9] shows another approach to automatic gate control systems. Authors have used RFID and the GSM module for the implementation of the system. The RFID used is active RFID. This acts as the transmitter module and the receiver module will be the tag reader at the check-post. ATmega328 Arduino controller is used to implement the system. It also has a SIM800 GSM module to send messages to the user after the toll amount has been deducted from the customer's savings account.

Automatic Number Plate recognition mostly for Indian number plates is studied and implemented in [10]. An embedded system is used which recognizes the number plate of the vehicles automatically. The authors implemented the system using a Digital Signal Processor TMS320DM6437 which is improved for image and video processing applications. The system integrates a number of algorithms. Trait based number plate localization is used for detecting the number plate of the vehicle. Image

Scissoring is used to segment the characters and to recognize the characters, statistical feature extraction is used. These algorithms were designed and used, taking in mind the conditions of the Indian license plates which vary widely in parameters such as font, size, color, position, etc. The system achieved a success rate of 82%. Accuracy of plate localization was found to be 87%, character recognition was 85% and character segmentation was 95%.

This paper presents a system that attempts to reduce the overall cost without affecting the efficiency of the system. It is well suited for small organizations. The main two technologies that are used are RFID for both pedestrians and vehicle users and authentication using license plate recognition using image processing techniques.

The rest of the paper is organized as follows: Sect. 2 presents the overview of the system. Section 3 presents the implementation followed by Sect. 4 which gives insights into the results. In the last section, the conclusion is discussed.

## 2 Overview of the Proposed Gate Control System

Figure 1 displays the overview of the proposed system. The management system is the main module that coordinates the proper functioning and communication among the different modules and the server. The pedestrian authentication module is responsible for verifying the identity of the pedestrians that intend to enter the premises of the

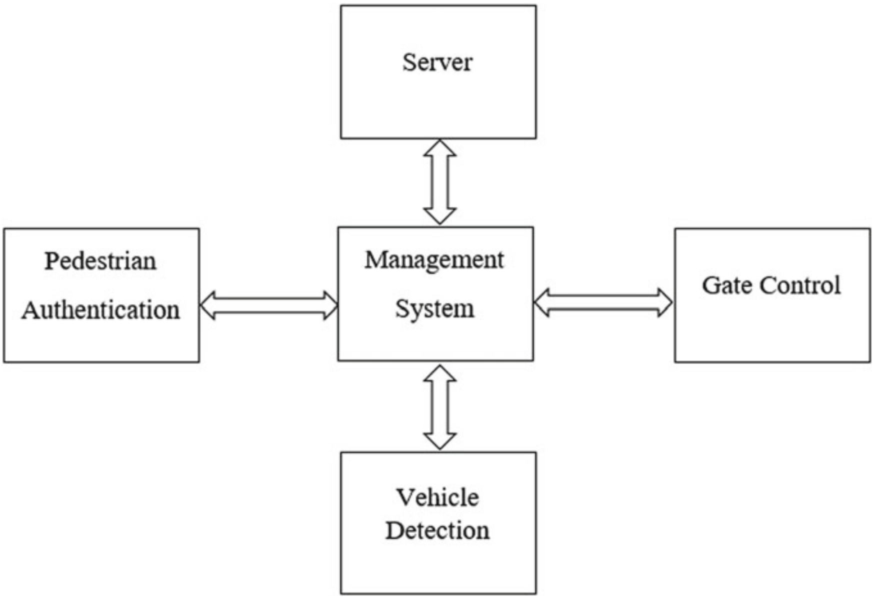



**Fig. 1** System overview

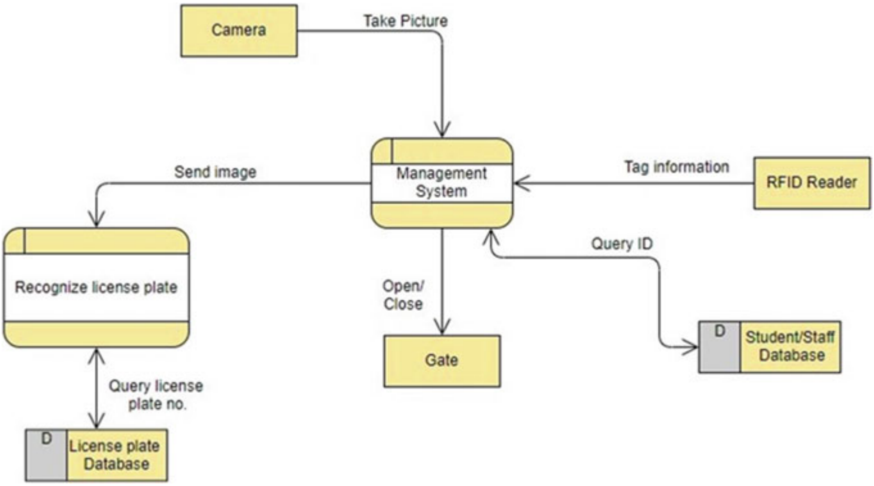



**Fig. 2** Data flow diagram

organization. The gate control module is responsible for the functioning of the gate. The management system decides when to open and close the gate. There is a camera and vehicle detection module that is in charge of the authentication of the vehicle. The image from the camera is taken which focuses on the number plate of the vehicle. This picture is used to extricate the number plate of the vehicle. The algorithm to extract the number plate is deployed on the server. If a match is discovered against the database for the number, then the gate opens. The Dataflow diagram is shown in Fig. 2. It tells about how the data flows in a system and its interaction. If the visitor is in a vehicle, then the image taken is sent to the cloud where license plate number is extracted and checked against the database. Gate is opened if a match is discovered. The other way to authenticate the visitor is when they swipe their ID card and the tag data is read by the RFID reader and checked against a database. Figure 3 represents the flowchart of the system. Vehicle is detected and then a snap is captured. The licensed number plate of the vehicle is obtained and then it is examined with the data archive. If there are any matches, then the gate is opened. If there are no matches, then ID is scanned. If there is an entry for the visitor in the database then the gate is opened, otherwise, access is denied.

## *2.1 Hardware*

The components utilized for this set up are shown in Fig. 4.

1. Raspberry pi: A small hand sized computer that helps developers and hobbyists create a prototype for hardware systems or use it as a general computer. It has several GPIO pins that could be utilized to work with a wide variety of sensors.

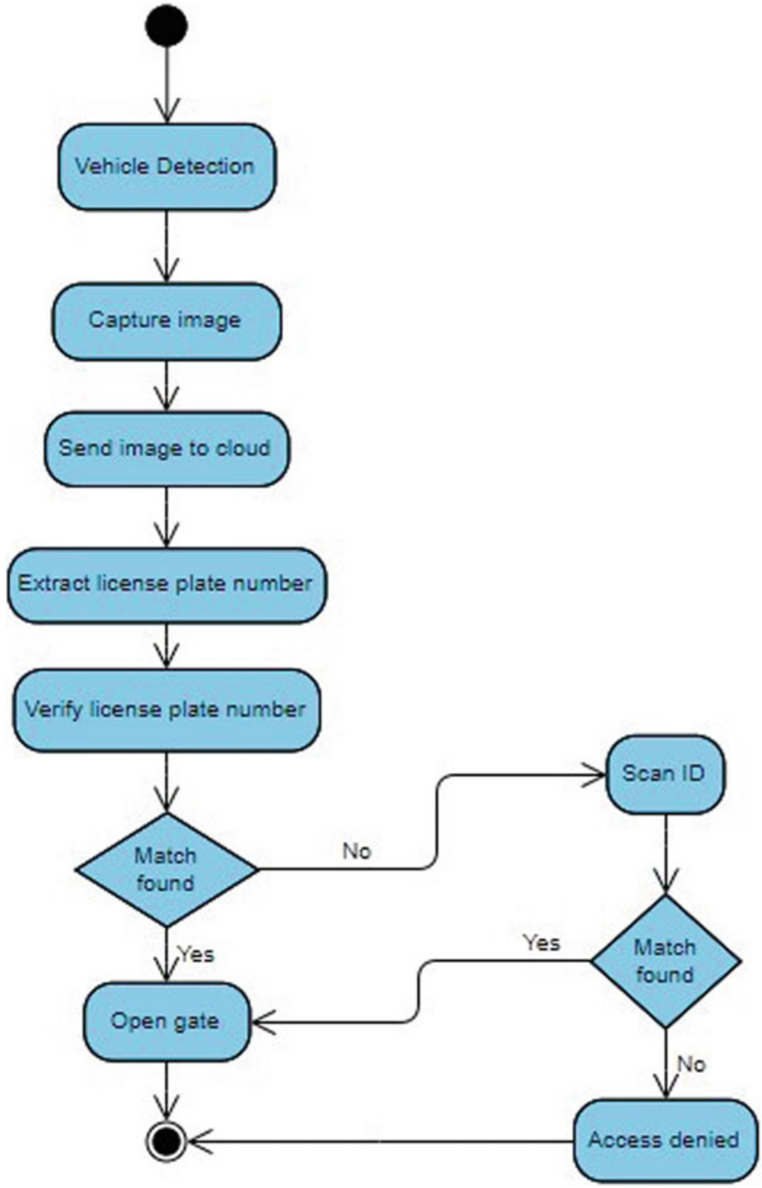
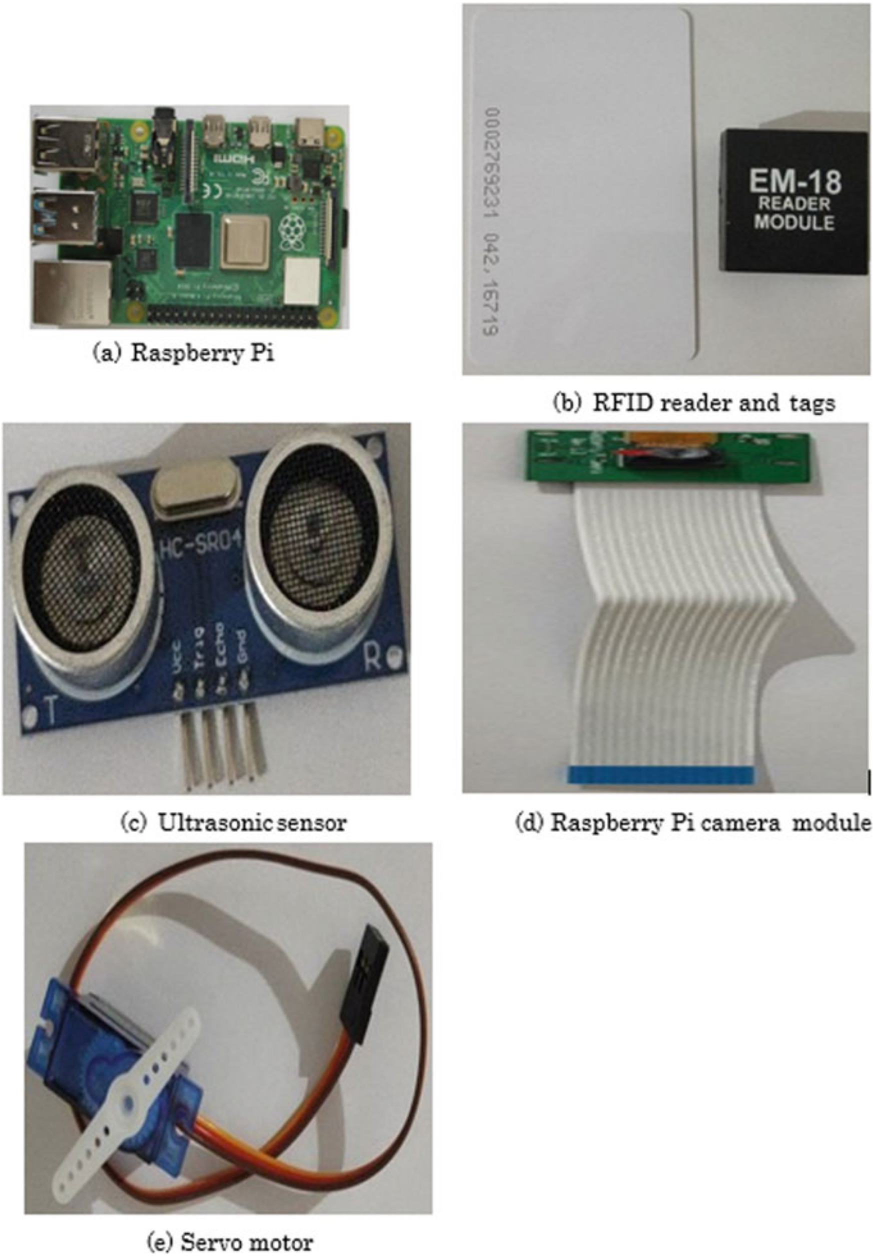



**Fig. 3** Flowchart of the system




(a) Raspberry Pi
(b) RFID reader and tags
(c) Ultrasonic sensor
(d) Raspberry Pi camera module
(e) Servo motor

**Fig. 4** Components

The model used for the implementation of the project is Raspberry pi 4. It is the primary management device.

2. RFID: RFID tags are utilized for authentication by making use of radio ways. To facilitate this an RFID interpreter is prescribed. The reader used is EM18.
3. Ultrasonic sensors: These are used to compute the proximity or distance to an object with the assistance of ultrasonic waves. The HC-SR04 ultrasonic sensor is used in the system to determine the existence of a vehicle. The range of the sensor is typically from 2 to 400 cm.
4. Camera: Raspberry pi camera module is used to capture the image of the incoming vehicle so as to extract the image of the region which focuses on the licensed plate number.
5. Servo motor: Servo motors are electronic devices used to rotate objects at specific angles. The SG90 9 g servo motor is used to operate the gate.

### *2.2 License Plate Number Extraction*

The computerized automated number plate acknowledgment can be mainly divided into 3 phases:

1. License plate detection: In this phase, the precise location of the number plate of the given vehicle is inferred. Localization of the license plate depends on the characteristics of the license plate. This stage is also exceedingly dependent on the country in which the system will be used. Each country has a different type of the licensed number plate. They usually differ in fonts, colors, and sizes. Based on the characteristics of the license plate potential candidate license plates are generated by following the given below steps:

   (a) Converting to Gray scale
   (b) Binarization of the image
   (c) Applying morphological operations

   Now from this list, a license plate is obtained by using the aspect ratio. In India, the standard size for the number plate is $200 \times 100$ mm for two and three wheelers, $340 \times 200$ mm for LMV, and $500 \times 120$ mm for passenger cars.
2. Character Segmentation: In this phase, the license plate detected from the previous phase is cropped out. The characters in the license plate are then separated individually to be used for recognition in the next phase.
3. Character Recognition: During this phase, the segmented characters are recognized. The last two phases in this project are taken care of by Pytesseract OCR. It is an Optical Character Recognition engine that is used to detect text which is embedded in images.

## 3 System Implementation

The main management system functions with the assistance of the Raspberry Pi. It is responsible for the overall working of the system. MySQL database is used to store the details of the vehicles and the pedestrians.

For pedestrians, RFID tags are given which are embedded in the ID card. An RFID interpreter is used to interpret the tags. The information so read is queried against a database. The pedestrian is allowed to enter the premises only if a match is discovered. The tags hold data that is stored electronically. The electromagnetic energy that is relayed by the interpreter powers the RFID tag as it lacks a dedicated power source.

Vehicles can be authenticated using the same RFID reader system. However, for better convenience, the ANPR system is also used. When a vehicle is advanced towards the gate, its presence is detected by an ultrasonic sensor. This triggers the snap of the vehicle to be captured by the Rpi system. The camera is positioned such that the focus is on the license plate. The seized image is sent to a server. This server is set up with the assistance of Google Cloud Compute Engine. Compute Engine is Google's IaaS (Infrastructure as a Service) which enables users to initiate Virtual Machines (VM) on demand. The ALPR algorithm deployed on the server receives the image and applies image processing techniques. The extracted number plate is run through Pytesseract, an Optical Character Recognition (OCR) engine which helps in extracting the number. The extracted number is queried against the MySQL database which is also hosted on the Google Cloud platform with the assistance of Cloud SQL. Cloud SQL is a completely managed database service from Google that helps set up and manage MySQL or PostgreSQL or SQL server.

The frontend of the system consists of a web application that will be used as a management system. It will be used mainly to register the details of a new vehicle or pedestrian and to monitor when someone tries to enter the premises. Additionally, there will be a facility to override access in case due to some technical reasons a genuine user is not authenticated. In such cases, the security personnel responsible for monitoring the functioning of the system can run a search against the database using the licensed number plate or the unique ID of the pedestrian. The web application is hosted online with the assistance of 000webhost which is a webhosting service. The UI is shown in Fig. 5.

## 4 Results

After running the algorithm to extract the license plate numbers it was found that for images taken from a straight line of sight, ROI (Region of Interest) was correctly detected. The first step that is localizing the license plate greatly depends on the camera angle. The OCR engine used was Tesseract-OCR. The engine was able to correctly extract the characters from the license plate. In some cases, some erroneous

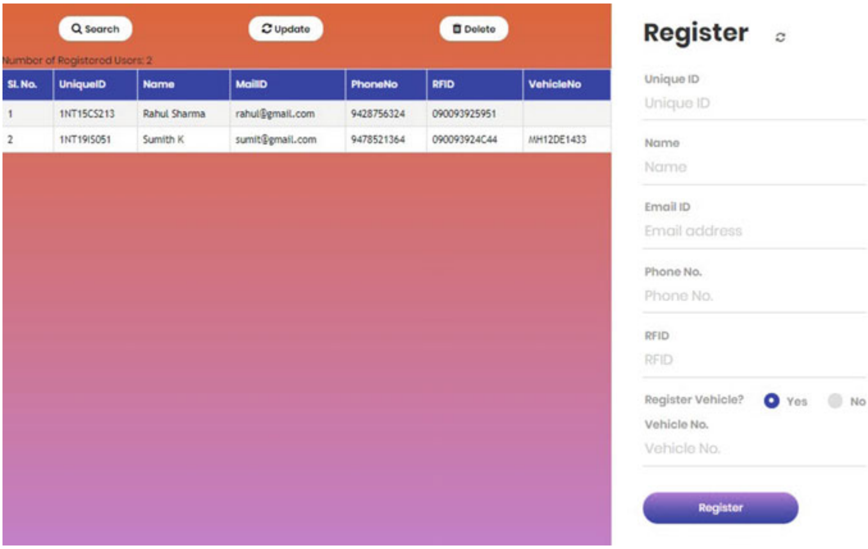
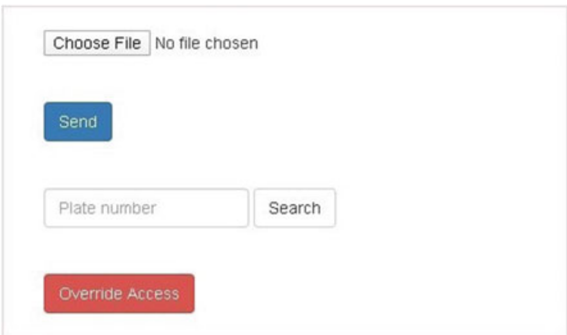



**Fig. 5** User interface of the management system

characters were also generated which was stripped off. Figure 6 shows the GUI used for testing the algorithm. The functioning part has an option to upload the file. The send button sends the image to the python file which in turn extracts the license plate number. The extracted license plate number will be shown in the textbox. The efficiency at which the number plate could be extracted greatly depends on how well the license plate can be localized from the image. This in turn depends on image processing algorithm used and the factors on which the localization depends. Object Detection using Haar feature-based cascade classifiers is a functional object noting method using simple features proposed by Paul Viola and Michael Jones in their paper [11], simple features like edge features, line features, four-rectangle features, and so on. It is an ML based approach where a cascade task edified from plenty of true and false images is later used to discover entities in supplemental images. On using

**Fig. 6** GUI used for testing



Object Detection using Haar feature-based cascade classifiers it was found that for few images the license plate region was being localized while the current algorithm failed to do so. Yet for some other images, the current algorithm could identify the license plate where the Haar Cascade Classifier failed to do so. An illustration of this can be shown in Figs. 7 and 8.

.


**Fig. 7** License plate is not detected using the algorithm

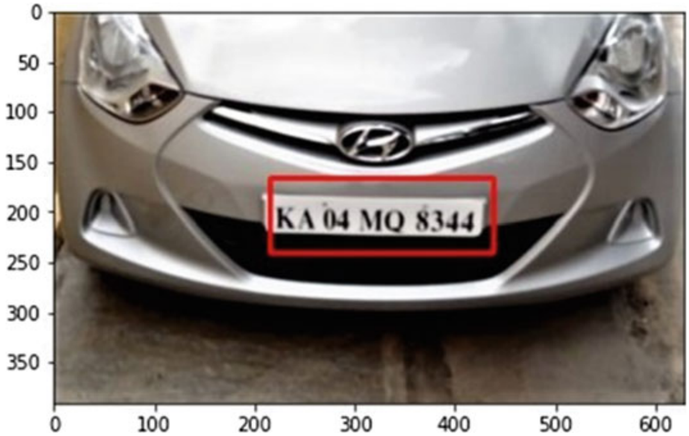



**Fig. 8** License plate is detected using Haar cascade classifier

## 5 Conclusion

ALPR (Automatic License Plate Recognition) defined as the automatic identification of the number plate mainly used for security systems and traffic monitoring systems. RFID is another technology that is used to verify visitor identity for security gate systems. The technology that is chosen to implement an automatic gate control system depends mainly on the cost, security, and efficiency of the system. The proposed system aims to implement a system that is cost effective and secure. This will help to ease the burden of the security personnel who have to deal with a lot of traffic especially during peak hours by handling the authorization of the visitors. The RFID system to authenticate pedestrians was found to be quick and effective. The image processing method of working is used to extricate the number plate area from the snapshots were not always effective. An improvement over this would be to use machine learning and build a model that can efficiently localize the license plate.

## References

1. Shan D, Ibrahim M, Shehata M, Badawy W (2012) Automatic license plate recognition (alpr): a state-of-the-art review. IEEE Trans Circuits Syst Video Technol 23(2):311–325
2. Hill JM, Cameron B (2000) Automatic identification and data collection: scanning into the future. Achieving Supply Chain Excellence Technol 7:269–272
3. Gunawan H, Kadir EA (2017) Integration protocol student academic information to campus RFID gate pass system. In: 4th International conference on electrical engineering, computer science and informatics (EECSI). IEEE, pp 1–6
4. Mahmood Z, Ali T, Khattak S, Khan SU, Yang LT (2015) Automatic vehicle detection and driver identification framework for secure vehicle parking. In: 2015 13th International conference on frontiers of information technology (FIT). IEEE, pp 6–11
5. Bani-Hani RM, Wahsheh YA, Al-Sarhan MB (2014) Secure QR code system. In: 10th International conference on innovations in information technology (IIT). IEEE, pp 1–6
6. Avinash M, Kishore Kumar R, Munusamy M, Muthuvel D, Naveen Kumar K, Thirumalaivasan K (2017) Multi-level security using QR code, biometrics and password for security systems. Int J Eng Sci 6334
7. Anawar S, Bahaman N, Ayob Z, Abdul A (2018) Implementation of intelligent automated gate system with QR code. System 9(10)
8. Kasym K, Sarsenen A, Segizbayev Z, Junuskaliyeva D, Ali MH (2018) Parking gate control based on mobile application. In: Joint 7th international conference on informatics, electronics and vision (ICIEV) and 2nd International conference on imaging, vision and pattern recognition (icIVPR). IEEE, pp 399–403
9. Balamurugan K, Elangovan S, Mahalakshmi R, Pavithra R (2017) Automatic check-post and fast track toll system using RFID and GSM module with security system. In: International conference on advances in electrical technology for green energy (ICAETGT). IEEE, pp 83–87
10. Kulkarni P, Khatri A, Banga P, Shah K (2009) Automatic number plate recognition (ANPR) system for Indian conditions. In: 19th International conference radioelektronika. IEEE, pp 111–114
11. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR, vol 1, pp I–I

# Smart College Camera Security System Using IOT

**Junaid, Mohammad Khalid, Namita Saunshi, Partha Mehta, and M. N. Thippeswamy**

## 1 Introduction

Conventional security systems have been traditionally basically manual. Manpower placed as security guards used to be all the security anyone needed to have in the days long gone. Later, as electronic cameras were developed, the trend changed to employ fewer security guards and have CCTV cameras installed at all strategic locations on campus. In case of any theft, the CCTV record could be reviewed by human operators to detect who the suspect could be [1] still, the system remained dependent on manual recognition of the intruders.

Gradually, with the development of image recognition and computer vision, more and more avenues opened up. All possible faces expected at a workplace could be fed into the master database into user wise folders and the newly captured random images could be labeled accordingly. This would do away with the necessity of having a human label a bunch of images and enabled fast tracking.

Employees who were registered in the database could be recognized and a log can be created for each employee with the time of check in and check out. Similarly, in educational institutes, individual logs can be created for students. More importantly, if a face is detected which is NOT in the database, it could trigger an alert or alarm and this will enable the security of the premises to be ensured [1, 2].

Raspberry pi is an innovative, powerful and small size computer used by most security systems. Security system built using Raspberry pi computing engine is more

Junaid · M. Khalid · N. Saunshi · M. N. Thippeswamy (✉)
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore, India
e-mail: thippeswamy.mn@nmit.ac.in

P. Mehta (✉)
PWC AC, Bangalore, India

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_26

efficient in comparison with other security systems as the user can control resources by remotely accessing them [1].

Important application of Video Security systems includes maintenance of discipline in prisons, provide personal and corporate security, and secure important moveable and non-moveable assets among others. As currently available video surveillance systems are pretty expensive a low-cost IOT based security system has been proposed [3].

Providing solutions to the above problem using images is an exciting and challenging task. The main objective of the research is to develop highly responsive cost effective security systems using Raspberry pi [4] as the basic computing element. It should be noted that the problem of facial recognition has been a subject of study for a long time and quite a few algorithms have been proposed for accomplishing this task. It should be noted that identification of persons based on facial image processing is a very difficult problem as human faces have a high degree of variation. Most of the earlier solutions were based on constructing a facial template and comparing it with the acquired facial image [5].

Identification was performed based on the closeness of match between the two. The significant features considered for the purpose of identification were the eyes, mouth, and nose. Using these facial components feature vectors were constructed. The mentioned approaches depend on the characterization and detection of the geometrical relationships between them and singular facial peculiarities if any. More recently algorithms which take a holistic towards face recognition without explicitly finding facial features have been proposed. A code is generated characterizing the face in an "n" dimensional space. This code is generated using various contortions applied to the target image in terms of viewpoint, scaling, aspect ratio, and wrinkling due to aging. Using all these features helps in building a robust detection scheme.

## *1.1 Contributions*

After studying conventional systems carefully, improvements in those systems were introduced with functionalities that conventional systems fail to provide. The flaws in the systems were compensated for by an intelligent security system. Following points explain the contributions in detail:

- A system is designed to capture the images or videos using a Raspberry Pi and Pi Camera. This particular subsystem captures the live stream and sends frame by frame to the receiver system that is requesting the live stream using IP and port number. The quality of frames can be improved by using better quality Pi camera, which is easily switchable. Frames are stored in buffer before they are sent to receiver to provide flow of control. The system provides live stream with very low latency.
- A face recognition algorithm has been implemented and a model for the network is trained using facial data. Usage of Pickles made the facial data encodings

persistent and repetitive training is not necessary for accurate face recognition. System provides near real-time face recognition where all faces which are known will be bounded by a box tagged with the name of the person (Unknown, if not identified). A record has been maintained for all the faces present currently in front of the camera.

- A web application has been developed to view the live video stream. The following three functions are defined in order to process the video and extract required information from it.
- Detection: Used to generate an alert when a particular person of interest (POI) comes into the FOV of the camera. The user enters the details of the POI. The system looks out for the particular person, and as soon as his presence is detected an alert will be generated. The system will subsequently keep tracking him and feed the relevant video to the user.
- History: In this mode, the system searches for the presence of the POI in the previously recorded video set, and forwards the same in chronological order to the user
- View Videos: In this mode, the selected recorded videos are played back to the user for manual analysis.

## 2 Proposed System

### *2.1 Design of the System*

In this section, several designs that are used in the process of designing the system are discussed. Figure 1 shows the design of the system, which shows a simple way as to how the sub components are working together. Intelligent cameras consist of a Raspberry Pi and a PiCam are digital cameras which capture the scene in real time and send the data for further processing.

There are two important sections, one is the Security System and other is the Intelligent Camera System. Intelligent camera will be using the python script to run the streaming and the security system will be executing python scripts to do several tasks such as showing livefeed, face recognition, live search, saving data, and systematic data search. The intelligent camera will be powered using a normal 2A adapter or an intermediate battery so that the system is impervious to power fluctuations. This use case diagram shown in Fig. 2 mainly has two actors: User and Service Provider.

1. **Service Provider**: The Service Provider is the one who installs and sets up the system. First, Service Provider will make sure the Security System and the Intelligent Camera are connected to the same network. Once that is done, the Scripts are executed on Raspberry Pi for starting the streaming. Secondly running the codes for grid view, live search and face recognized live feed.

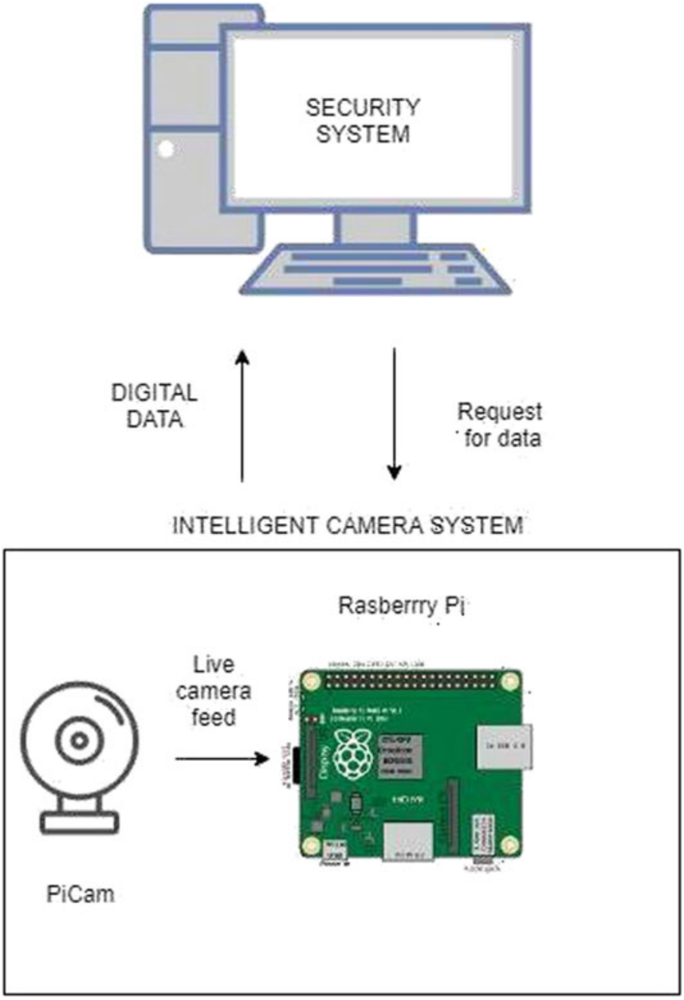



**Fig. 1** Architectural design of the system

2. **User**: The User can choose the function based on the processing requirements. The outputs will be displayed in a browser window. The choices are:
   (a) LiveFeed
   (b) Search in OldData
   (c) GridView
   (d) LiveSearch
   (e) Face Recognized Live Feed

The database of facial images of POI is prepared offline and provided by the service provider. The user selects his activity mainly- Live Feed, Live Search, Grid

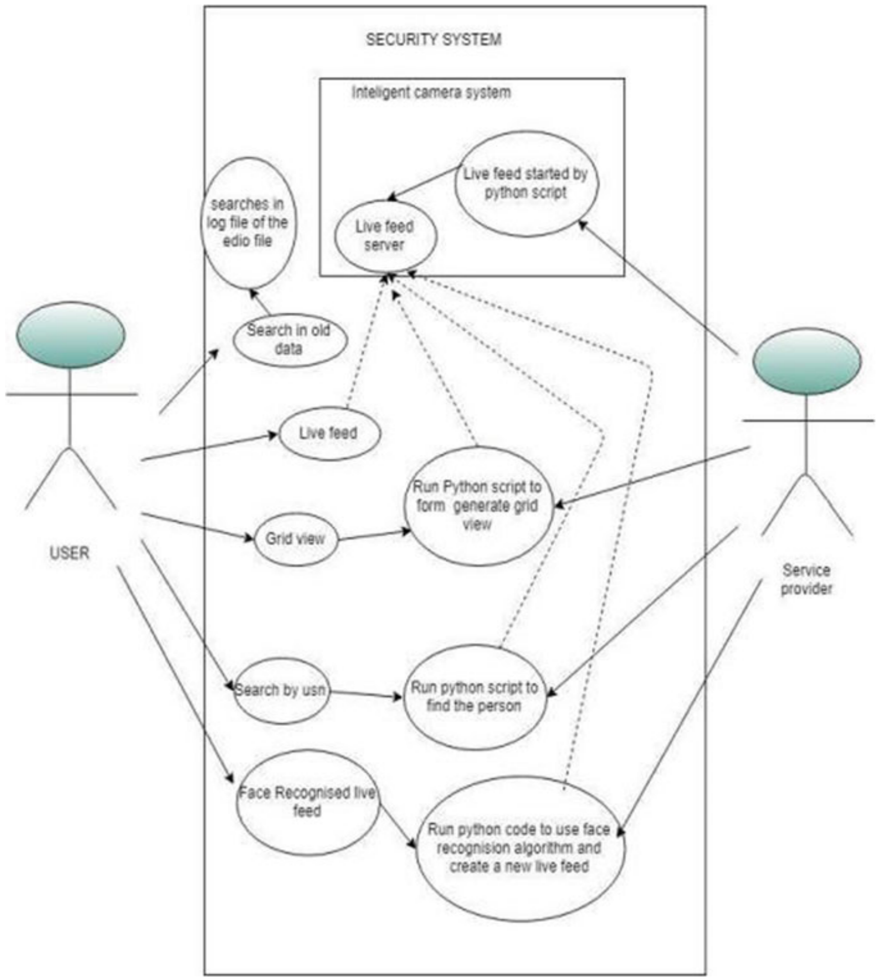



**Fig. 2** Usecase diagram of the security system

View, Search from Old Data, and Live Face Recognition. Live feed and searching through old data are the options which are mostly exercised.

The main Sequence of operation (refer to Fig. 3) in this system are:

1. The user will request for live feed and then the system will make a request to the intelligent camera system and get the digital data and will display the feed on the browser.
2. The user requests for the grid view and the system will request the intelligent camera for the live feed. Frames are processed and then a grid view is displayed on the browser.

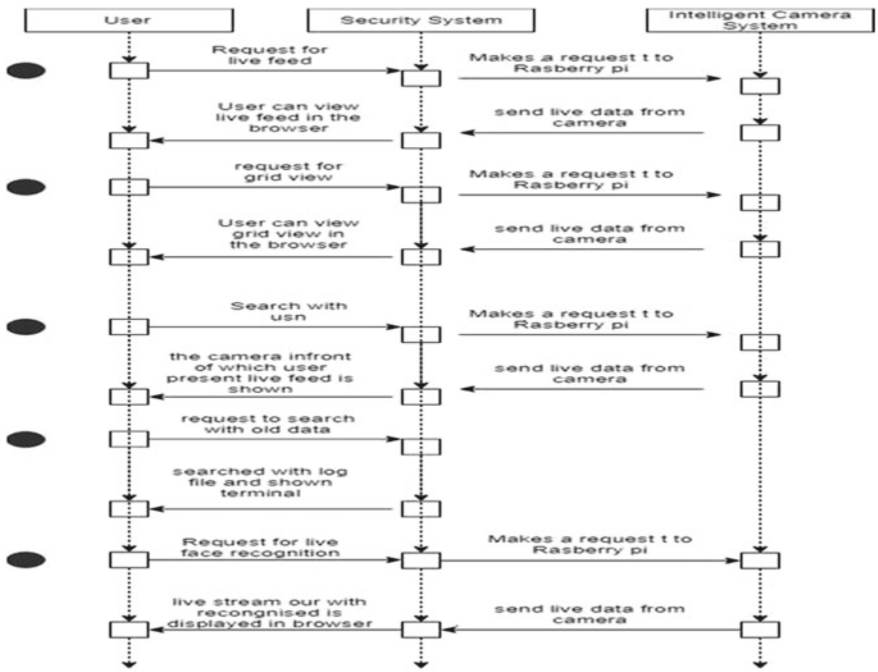


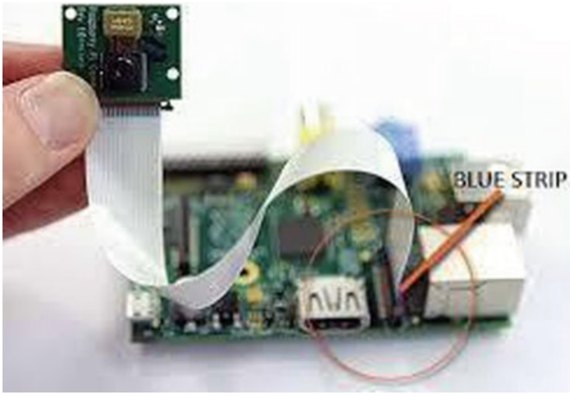

**Fig. 3** Sequential diagram of all sequences in the security system

3. The user will request for live search with USN. The user will enter the details and the system will make a request to the intelligent camera system to give the live feed and then it is processed for face recognition using an already trained model and then the USN is cross checked with the USN entered if it matches the camera live feed is shown to the user.
4. The User will request for searching the pre-recorded data and will enter the details of the POI. The system will systematically search through the recorded data. The matched records will be shown with details on the browser window.
5. The user will request for Live Face Recognized feed and the system will request the Intelligent Camera System to give live feed and then process the data with an already trained model to produce a near live stream of feed with recognized faces in the browser.

## 2.2 Methodology

Figure 4 shows the methodology and explains the method and the techniques which are used to implement the work. The camera module is connected to the raspberry pi through a strip which gives raspberry pi access to the camera feed. The inbuilt wifi module in raspberry pi provides the connection to the router. The client system

**Fig. 4** Shows the connection of the components in the system




will also be connected to the same router. As mentioned in the user requires all users have to be present in the same network. Since they are present in the same network, the data transaction is pretty fast. The user starts using the application by opening the browser. Four options are provided:

- Live Stream
- Display Detected People
- Live Face Recognition
- Search (by face and old data)

As per the options chosen by the user the raspberry pi sends live feed or the pre-recorded data to the user. Based on the data received the client system will process the information and will produce the required objectives.

## *2.3 Description of Process*

This section explains the internal working of each process used in the system and also discusses various algorithms and steps taken in the development of the process. It also describes the various applications of the system and how it is working.

1. **LiveStream:**
   Raspberry Pi acts as the server and sends the live stream (captured data by Pi Cam) in MJPEG format using sockets over HTTP. Each frame is sent as JPEG image. A buffer of frames is maintained from which frames are sent to client. Live stream can be viewed using IP and Port Number.
2. **HumanDetection:**
   Human detection is implemented using OpenCV (CV2 Library in Python3). Client receives frames and performs detection on every image with the help of pretrained detector. OpenCV human detection uses HOG (Histogram of Oriented Gradient) descriptor to select an appropriate region in every image.

SVM (Support Vector Machine) classifier runs on the selected region and object is classified as human. Each person detected is converted to individual image and is stored in Python Flask Gallery.

3. **Face Recognition**
   Face recognition is performed on the live stream data. This module puts the classification label on a known face with a respective name and USN using: (a) OpenCV (b) Deep learning (c) Python. Live Stream is fed as input to the python file as shown in Fig. 5.2. Necessary Python packages for face recognition are dlib and face recognition libraries. These libraries are open-source and provide an easier way to implement machine learning algorithms and to build an effective face recognition system. A data set is maintained for every person with multiple images per person. Face encodings are generated using 128-d vectors per image [6]. A pickle file is maintained that has face encodings for every person. Encodings are stored as key-value pairs with key as name of the person and values as a list of encoded vectors [6]. Pickle file acts as the trained model [6]. The recognizer program takes live stream as input, detects faces in every frame, and every detected frame and encoding of the detected face is compared with all encodings in the pickle file. Euclidean distance is calculated and if this distance is below a certain threshold value, it is considered as a vote. The face with maximum number of votes is considered as the output classification label [7].
   In the background, every frame is recorded in a video file that can be used as a future reference to provide better surveillance or observation. Simultaneously, track of all faces at any instance is stored in a temporary file. A search option is provided to search by a unique id. If the search is successful, page will be redirected to the camera live stream where the person is present [7] otherwise, outputs an unsuccessful search. A log file is also maintained to keep track of people present in front of the camera along with the time stamp. It clearly tells on what date and at what time the person was present in front of the camera [7].

## 3 Results

In this section, the results which are obtained after testing the system are discussed. Mainly, how the user interface is developed along with some test data and how the results are produced.

Figure 5 shows the user the Home Page where options for either to upload Name and USN, search option and either face search or log search. Face search is to search for the faces of different students or any other person, log search is to search students present in front of camera depending on the time stamp and data. At last, there is an option to go to live streaming.

Figure 6 shows how to create a folder with student's name. Entering the Name and USN of the student will lead to creating a new folder for adding the required images to the folder.


**Fig. 5** Home page

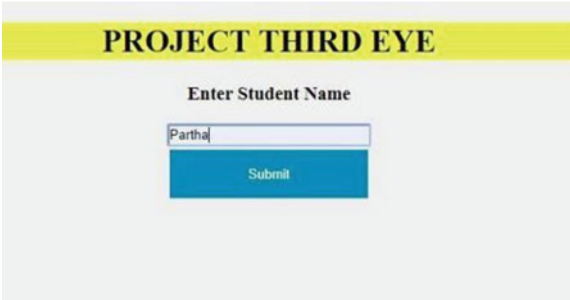



**Fig. 6** Creating data set of new students

Figure 7 shows that by clicking on the "Choose Files", a file browser is shown, which is capable of selecting multiple files to be uploaded to created folder with student's name.

Figure 8 shows that it gives you options of search such as Face Search, Individual Search, and Display All Record. Face Search provides you search according to the

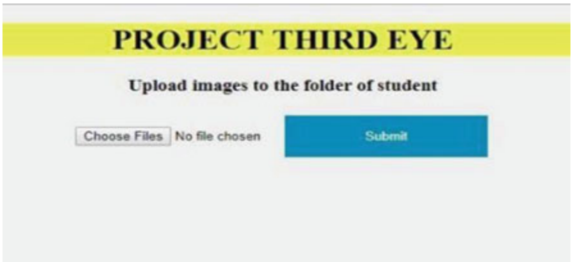



**Fig. 7** Browsing multiple files to be uploaded


**Fig. 8** Home page of Searches


**Fig. 9** Live searching of face


**Fig. 10** Individual search

faces recognized of a particular person and is found in detected people database or not. Clicking on face search takes you to live streaming video. Individual record gives you an option to search for a particular person's details according to time stamp. Display all records: gives you an option to display the user's data in the logfile.

Figure 9 shows that here by entering the name of the person will be able to do a Live Search and point to live stream which is being taken from Raspberry Pi and Pi camera as shown in Fig. 10 or not found case in Fig. 10. Figure 10 shows that student is not found in the detected people's list.

Figure 10 shows that the search for the presence of a particular person in the college from the time when the program started. It will find a person in particular according to date and time as to when and where the person was in front of camera.

Figure 11 shows that face recognition of this individual person will be printed along with the time stamp. Here the particular person's whereabouts are displayed (Junaid in this case).

Figure 12 shows that the whole record of recognition is displayed on the browser. Contains details of all known and unknown people. This will just print all the students who are detected in the camera and present them in a log file.

Figure 13 shows the live feed button. When it is clicked the page redirect to live stream page which contains the live feed from the camera.

Figure 14 shows the live stream which is used to produce a grid view page in such a way that the grids will have all the humans detected as separate images. Continuous updates will cause the browser to auto reload and validate itself to show the updates in the browser. In the next application, face is detected and classified based upon the trained images (training data). Figure 15 shows the face (Known Face) model. Person present in front of the camera is recognized in near real time. So, the face will be known if that person's or student's face is trained at the time of training. Figure 16 shows the next application and provides the live feed to the face recognition. The subject is unknown to the system due to the absence of facial data in the data set.

**Project Third Eye**

Junaid-15CS068 : 2019_03_30_03_03_29
Junaid-15CS068 : 2019_03_30_03_03_30
Junaid-15CS068 : 2019_03_30_03_03_31
Junaid-15CS068 : 2019_03_30_03_04_48
Junaid-15CS068 : 2019_03_30_03_04_49
Junaid-15CS068 : 2019_03_30_03_04_50
Junaid-15CS068 : 2019_03_30_03_04_51
Junaid-15CS068 : 2019_03_31_16_25_35
Junaid-15CS068 : 2019_04_01_01_42_23
Junaid-15CS068 : 2019_04_01_01_42_26
Junaid-15CS068 : 2019_04_01_01_42_27
Junaid-15CS068 : 2019_04_01_01_52_26
Junaid-15CS068 : 2019_04_01_01_52_28
Junaid-15CS068 : 2019_04_01_02_14_46
Junaid-15CS068 : 2019_04_01_02_40_19
Junaid-15CS068 : 2019_04_01_12:52

**Fig. 11** Individual search result

Figure 16 shows unknown face. The face is recognized as unknown for a person who is standing in front of the camera.

Figure 16 shows the test case for known face recognition. The live feed to the face recognition algorithm which has multiple faces that are known, standing in front of the camera. In this case, multiple students can be recognized showing their names and USNs around the bounding box. Figure 16 shows the live feed produced by the face recognition algorithm will have a bounding box around the faces with Name and USN of the person labeled outside the box. Known faces have names and USN's as classification labels (Fig. 17).

## 4 Conclusions

There is a huge requirement for better security systems and this system is expected to satisfy the desired need of the user. Conventional security systems are not satisfying user needs because they lack high level understanding of the digital data that are received. Smart College Camera Security system uses smart cameras that is capable of drawing a real time actionable insight using image data. Smart College Camera Security system need not be monitored by multiple security personnel as it is capable of taking decision on its own and keeps track of all individuals that are captured. Further, this system is comparatively less costly than a system manufactured by tech giants. This system can be easily installed in colleges at strategic locations. Further,

**Project Third Eye**

Khalid-15CS095 : 2019_03_30_03_02_22
Khalid-15CS095 : 2019_03_30_03_02_23
Khalid-15CS095 : 2019_03_30_03_02_24
Khalid-15CS095 : 2019_03_30_03_02_25
Khalid-15CS095 : 2019_03_30_03_02_26
Khalid-15CS095 : 2019_03_30_03_02_27
Khalid-15CS095 : 2019_03_30_03_02_28
Khalid-15CS095 : 2019_03_30_03_02_29
Khalid-15CS095 : 2019_03_30_03_03_29
Junaid-15CS068 : 2019_03_30_03_03_29
Khalid-15CS095 : 2019_03_30_03_03_30
Junaid-15CS068 : 2019_03_30_03_03_30
Junaid-15CS068 : 2019_03_30_03_03_31
Khalid-15CS095 : 2019_03_30_03_03_31
Khalid-15CS095 : 2019_03_30_03_03_32
Khalid-15CS095 : 2019_03_30_03_03_33
Khalid-15CS095 : 2019_03_30_03_03_34
Khalid-15CS095 : 2019_03_30_03_03_35
Khalid-15CS095 : 2019_03_30_03_03_36
Khalid-15CS095 : 2019_03_30_03_04_47
Khalid-15CS095 : 2019_03_30_03_04_48
Junaid-15CS068 : 2019_03_30_03_04_48
Junaid-15CS068 : 2019_03_30_03_04_49
Junaid-15CS068 : 2019_03_30_03_04_50
Khalid-15CS095 : 2019_03_30_03_04_51
Junaid-15CS068 : 2019_03_30_03_04_51
Khalid-15CS095 : 2019_03_30_03_14_27
Unknown : 2019_03_30_03_14_27
Khalid-15CS095 : 2019_03_30_03_14_28
Unknown : 2019_03_30_03_14_28
Unknown : 2019_03_30_03_14_29

**Fig. 12** Whole systematic record output

**Fig. 13** Live feed

**Fig. 14** Test image for human detection

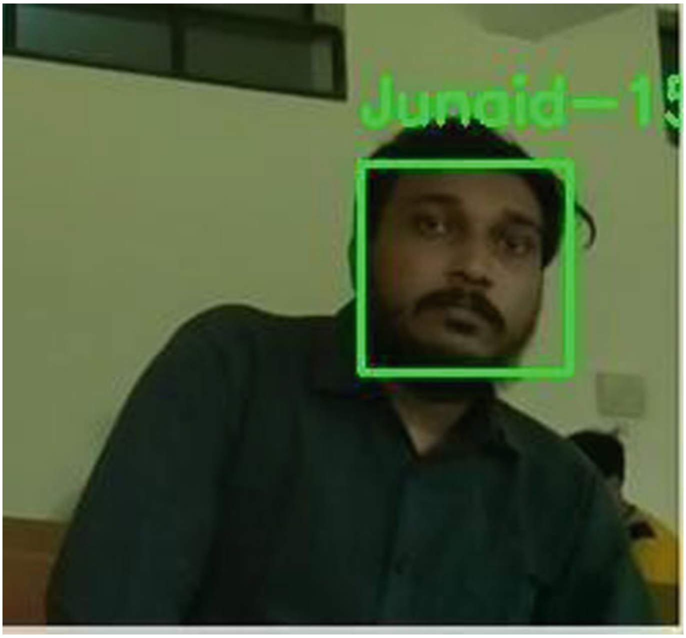



**Fig. 15** Live face recognition (Known Face)

these systems provide better results and accuracy than conventional camera security systems. These systems are capable of capturing any sort of minute detail that can be missed by an ordinary security camera.

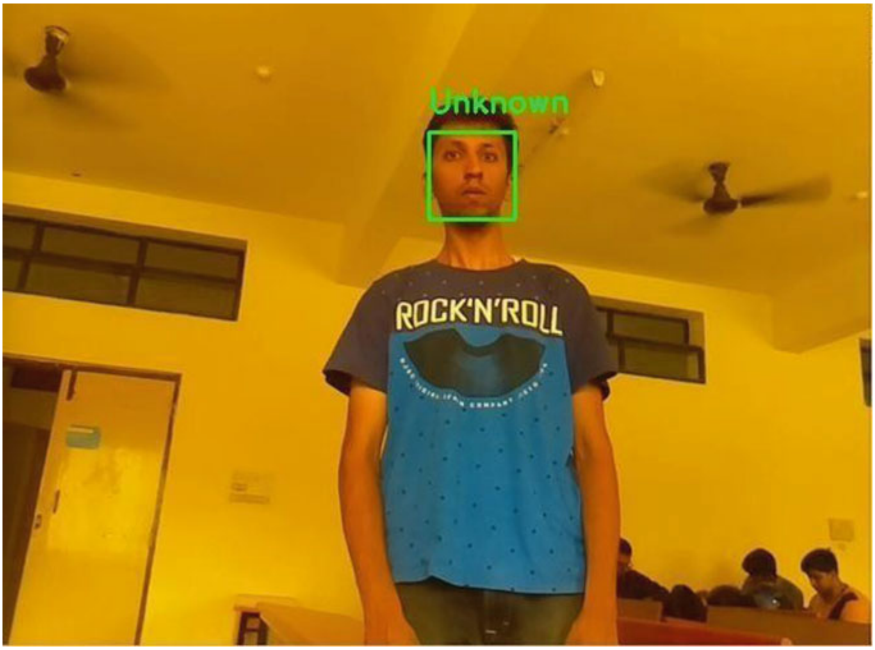



**Fig. 16** Live face recognition (Unknown Face)


**Fig. 17** Live face recognition multiple (Known Face)

## References

1. Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. IEEE Trans Pattern Anal Mach Intell 28(12):2037–2041. https://doi.org/10.1109/TPAMI.2006.244
2. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19(7):711–720. https://doi.org/10.1109/34.598228
3. Belongie S, Malik J, Puzicha J (2002) Shape matching and object recognition using shape contexts. IEEE Trans Pattern Anal Mach Intell 24(4):509–522. https://doi.org/10.1109/34.993558
4. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE conference on computer vision and pattern recognition, columbus, OH, pp 580–587. https://doi.org/10.1109/CVPR.2014.81
5. Guodong C, Xia Z, Sun R, Wang Z, Ren Z, Sun, L (2011) A learning algorithm for model based object detection. In: 2011 8th international conference on ubiquitous robots and ambient intelligence (URAI), Incheon, pp 101–106. https://doi.org/10.1109/URAI.2011.6145941
6. Liu S, Deng W (2015) Very deep convolutional neural network based image classification using small training sample size. In: 2015 3rd IAPR Asian conference on pattern recognition (ACPR), Kuala Lumpur, pp 730–734. https://doi.org/10.1109/ACPR.2015.7486599
7. Punjabi A, Katsaggelos AK (2017) Visualization of feature evolution during convolutional neural network training. In: 2017 25th European signal processing conference (EUSIPCO), Kos, pp 311–315.https://doi.org/10.23919/EUSIPCO.2017.8081219

# Aquatic Debris Detection System

## Image Processing—IOT Integration Perspective

**Kubra Fathima, H. R. Preethi, Pinki, Rekha Myali, and N. Nalini**

## 1 Introduction

Aquatic debris—is found in aquatic environments, such as streams, rivers, ponds and lakes. Debris can be defined as any persistent solid material or waste created by humans that is found in water bodies that has emerged to be a serious environmental issue. Marine debris is an important environmental issue [8]. It is a threat to our ecosystems, water transport and life of living things in both land and water. So monitoring aquatic debris is very important to the entire world. Water is the fundamental need for a number of activities, such as consumption, agriculture or tour as leads in conformity with pollution of lot of water bodies. Water sources are facing severe threats from floating debris. Due to the tremendous increase in globalization and industrialization, shifting of people from rural to urban areas and vast usage of land and water resources, the overall quality of water available to people has declined greatly. The vast utilization of fertilizers by farmers in agricultural fields including other chemicals in various sectors such as construction and mining have contributed immensely to the increase in debris, leading to deaths of aquatic life and overall reduction of water quality globally [1]. Organic and nutrient material loads cause severe water quality problems. More than 90% of the River Basin Management Plans (RBMP) assessed stated that agriculture is a significant pressure in the basin and also indicated that it is a main source of pollution by pesticides, organic matter, nutrients and hydro morphological impacts [2].

Consequently, there is a need to improve existing framework for checking water bodies and flotsam and jetsam, given that lab strategies are extremely delayed to

K. Fathima (✉) · H. R. Preethi · Pinki · R. Myali · N. Nalini
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology Bengaluru, Bengaluru, Karnataka, India

N. Nalini
e-mail: nalini.n@nmit.ac.in

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_27

build up an operational reaction and furthermore doesn't give general wellbeing security on ongoing premise. Improved observing and evaluation instruments to build up a factually vigorous and extensive image of the general status of the oceanic condition with the end goal of further arranging. Nowadays Internet of Things (IoT), Image Processing, Remote Sensing (RS) and various other technologies are used in various area of research for monitoring, collecting and analyzing of data from remote locations [1]. The water quality monitoring is necessary which includes collecting data of several chemical parameters of water in real time and also capturing of image of water body if debris is present and transfers the collected data through a Wireless Sensor Network (WSN) for further analysis of data of water body. The rest of the paper is organized as follows:

- we analyzed the various reports and statistics related to water pollution levels and their impacts. Then, we reviewed various techniques used for developing debris detection and water quality monitoring systems in the past few decades and their shortcomings along with their advantages (Sect. 2).
- we review the overall architecture framework of the designed embedded platform used for sensing and analyzing the data. It provides complete guidelines regarding the inputs of the system, centralized control unit and outputs of the system (Sect. 3).
- we review the methodology and algorithms that are used for building the cost-efficient working framework of debris and water quality monitoring, along with the advantages of the model (Sect. 4).
- we put the data flow diagram (DFD) of the model, which represents the flow of data in the system (usually an information system). It is the graphical representation to visualize the flow of data in terms of inputs and outputs (Sect. 5).
- Final results obtained are explained along with snapshots (Sect. 6).
- Based on the developed model, we concluded the overall work done to implement the prototype of the model (Sect. 7).
- Based on the developed prototype, future scopes of the project are summarized (Sect. 8).

## 2 Literature Survey

In the last few years, debris monitoring was basically done by humans by manual spotting using watercrafts or patrol ships. But, this process requires more man power and is only applicable for small-scale monitoring of aquatic debris. Thereafter, by using advanced technologies and robotic platforms, autonomous underwater vehicles were used to sense the aquatic environment. Nevertheless, this also had some disadvantages namely high manufacturing costs, huge weight and large size restrict their deployment over large areas for monitoring surface debris objects on water. In the recent days, remote sensing is used as an alternative method for monitoring aquatic debris in large water bodies. However, this approach is effective only for short-term monitoring of highly concentrated aquatic environment. Moreover, sensing platforms with embedded sensors along with camera sensors provide more efficient and cost effective solution for observing and detecting debris in aquatic environments, so we focused on developing a multipurpose aquatic debris and water quality sensing platform for remote and long-term monitoring of water body. Scientists reported that there are 5.25 trillion pieces of plastic debris in the ocean. Of the mass 269,000 tons float on the surface, while some four billion plastic microfibers per square kilometre litter the deep sea. Figure 2 shows the graphical representation of pathways through which plastic enters the ocean bodies [3]. Recently, Ministry of Consumer Affairs, Food and Public Distribution has released Water Quality Report for State capitals and Delhi as analyzed by the Bureau of India Standards (BIS). Tap water in Mumbai is the safest for drinking while Delhi's water is one of the worst. In this advanced technology, there are some projects for water quality monitoring and debris detection. At initial stage, a low cost system for real time monitoring and assessment of potable water quality at consumer sites using various electrochemical and optical sensors was developed. The model was focused to be light-weight, cost effective and to use for long time monitoring. However, this implementation was highly recommended for large deployments using a sensor network approach for providing comprehensive rich data to water consumers, water companies and water authorities. A large-scale survey and market research is performed to identify low cost sensors that can accurately collect and monitor various parameters which can be used to get information about the overall quality of the water body [4]. Based on the parameters that are selected, an array of sensors is constructed that includes multiple microsystems which are used for analogue signal conditioning, processing and remote presentation of collected data. At last, an algorithm for fusing various sensor readings is developed to get details regarding water contamination risk. Thereafter, Nitrate and Sulphate Estimations in Water Sources Using a Planar Electromagnetic Sensor Array and Artificial Neural Network Method was proposed. The principle advantages of planar electromagnetic sensors are they are available for cheap, appropriate for on-site monitoring systems, instant reaction, and highly resistant. In further stages, wireless sensor networks was used for online monitoring of quality of water using several parameters such as turbidity, dissolved oxygen, conductivity, dissolved metal ions and pH levels in water body was analyzed and pollution levels in water was reported. It is estimated that future generation of wireless sensor networks can be successfully combined with robust, accurate, low cost and miniature sensors based on thin-film or thick-film semiconductor sensing electrode, which can reduce difficulties associated with bio-fouling [5] (Fig. 1).

## 3 Architectural Framework

The framework consist of aquatic sensor nodes for measuring quality of water and implement the debris detection algorithm using MATLAB. The realistic deployment shows our approach consumes less energy consumption and hardware resources [5]. Here we get data related to quality of water through various sensors like pH level

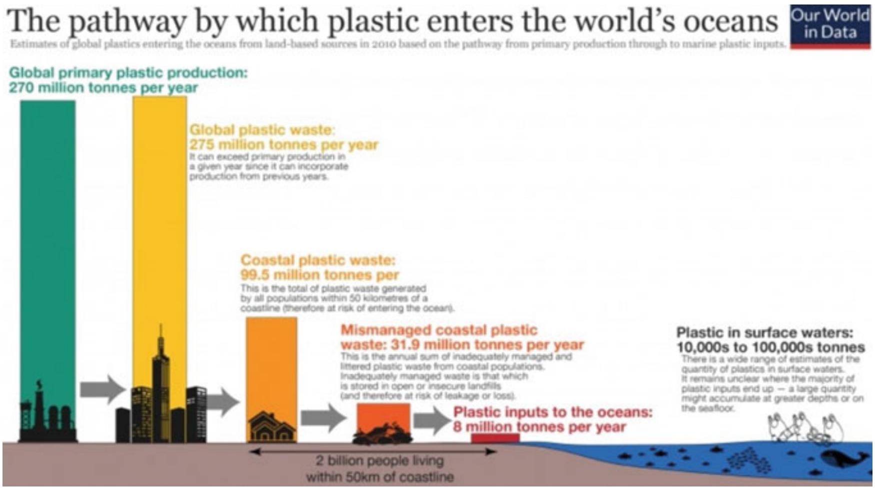



**Fig. 1** Plastic pollution in water bodies. *Source* Adapted from [3]

sensor to get information regarding acidity levels of water body, Dallas Temperature sensor which measures underwater temperature of the aquatic body and turbidity sensor is used to measure clearness of the water body. All the sensors are connected to the Microcontroller. The power to the entire system is provided using a batteries. The information will be uploaded continuously to cloud from the WSN through Microcontroller and Wi-Fi connected to laptop. We monitor this data and upload this data to cloud and end-users can obtain this data through Blynk App. Figure 3 represents the architecture design of the prototype developed to monitor the water body.

The debris images can be captured by any camera device present in any of the smart phones. When the debris are present on the water surface, the image of the water body is captured and is uploaded through WSN to the Matlab software. Furthermore, segmentation and analysis of image is done, then RGB colour image are converted grayscale image using algorithms and morphological operations algorithms are applied and further processing of the images is carried out. Once the complete analysis and processing is done, if the debris are detected on the water body, then an alert sound is produced to notify user and an email will be sent to the concerned water authority along with image of the water body captured saying" Urgent action is required…!! Clean the Water body…!!!". The developed frontend web application clearly reports the levels of debris in the water body and presence or absence of debris along with grayscale image and black and white image of the water body. Debris can be detected simultaneously in all subparts of the water body that is segmented. Initially the system needed to be trained then the testing can be done. The concentration values of the debris displayed is directly proportion to the dimensions of the image of the water body subjected for analysis.

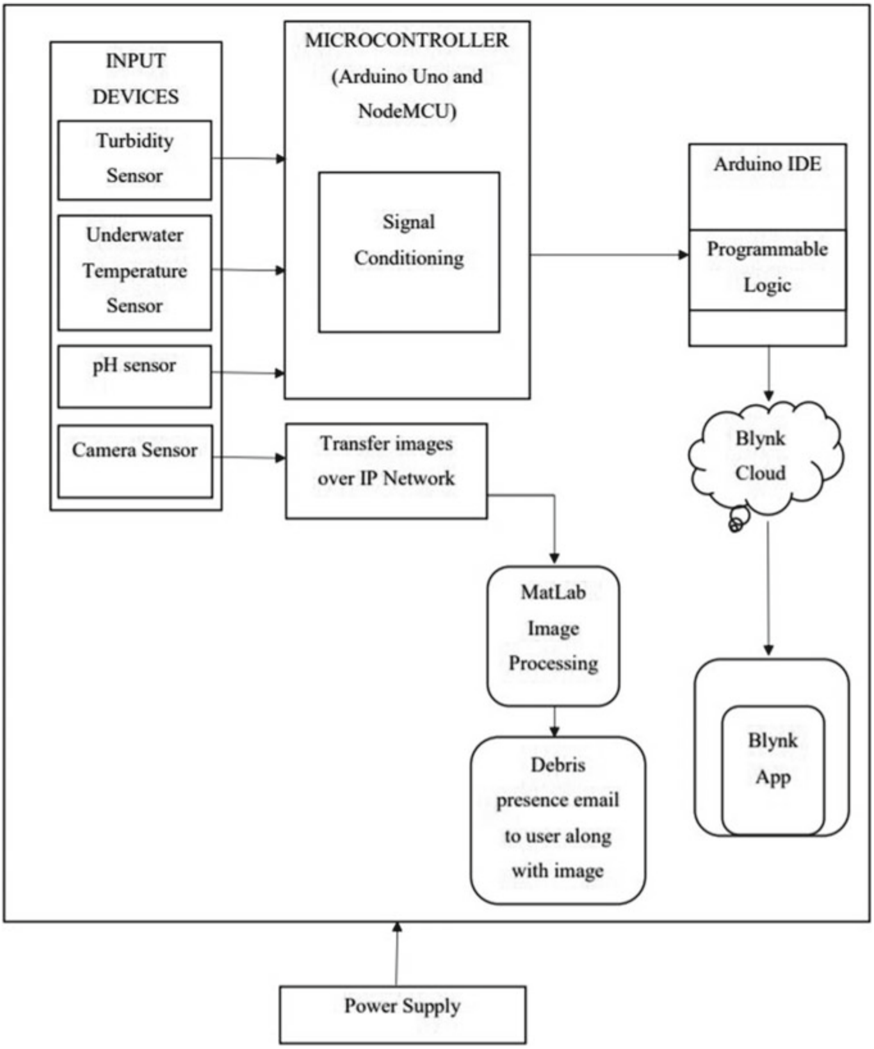



**Fig. 2** Architectural design

## 4 Methodology

In this advanced technology and innovations in the domain of image processing, sensing platforms, communication systems, wireless sensor networks, cloud computing and data analytics, it is possible to develop an integrated sensing system which can monitor both water quality and debris on the water bodies effectively. These advanced technologies have life-saving and environment-saving potential in these high pollution situations. Internet of things (IoT) is one such technology to build

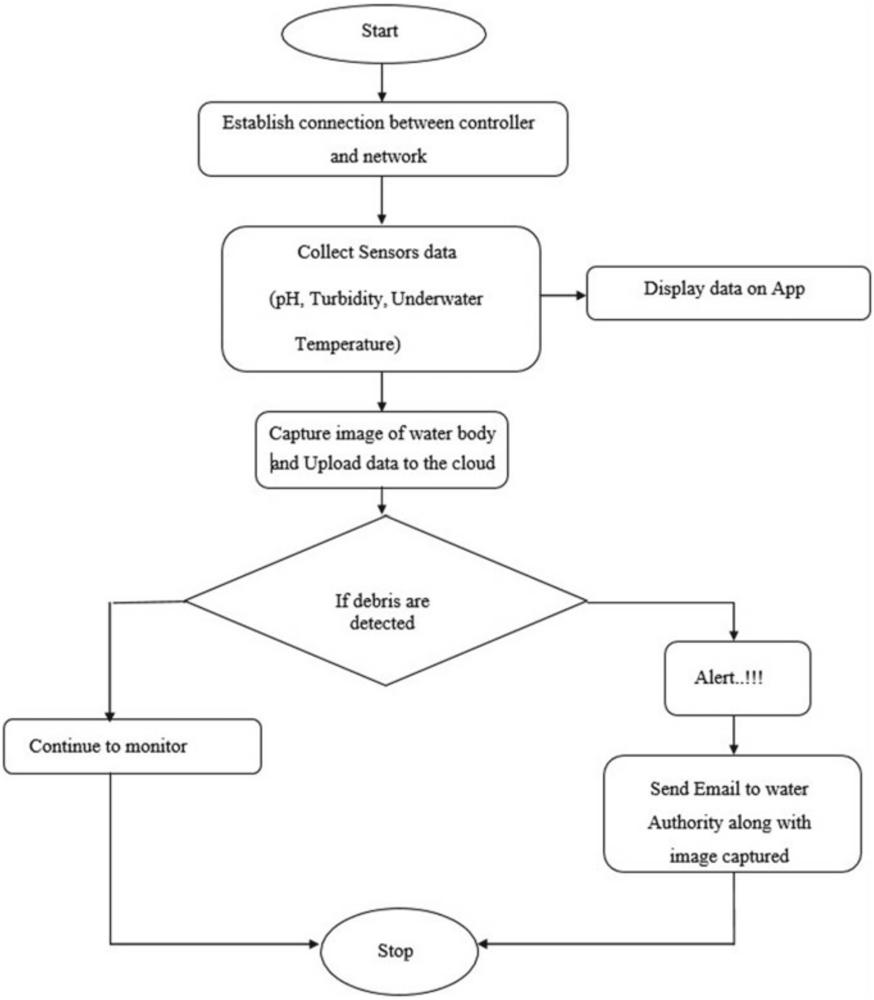



**Fig. 3** Data flow diagram

Smart projects and contribute to the growth of Indian economy. The Aquatic Debris Monitoring System is built using Internet of Things (IoT) and Image Processing. The model follows incremental and iterative approach. The ultimate goal is to build an integrated and embedded sensing platform to monitor the water quality along with detection of debris in water and develop a frontend to provide essential information to the user regarding levels of debris, an android Blynk app to provide necessary data regarding quality of water and an email service to alert the water authorities to clean the respective water body. The application also makes use of fuzzy logic algorithm for making decisions at correct time.

The embedded sensor board consist of various kinds of aquatic sensors namely pH sensor, turbidity sensor and underwater temperature sensor. The sensor node periodically collect the water quality data such as acidity level of water, transparency of water, etc. and the IP camera app captures images of the water body for debris detection. We propose an accurate and computationally efficient approach for acquisition, processing and transmission of water quality data as well as for detecting debris in the water body. On the basis of compressive sensing, our method allows the user to perform further segmentation on the captured image of water body, divide the image into subparts and detect debris simultaneously in all the subparts of the image by retaining most using information. More importantly, the algorithm achieves different granularities in detection of debris depending on various variations and dimensions of the aquatic environment. The system developed is environment-friendly, which doesn't harm any aquatic bodies or cause any disturbances to the aquatic life. The prototype is simple, portable, cost effective and also easy to maintain. We validate the efficiency in terms of detection reliability and average runtime through extensive experiments and comparisons. Be that as it may, there are some particular issues anticipating answers for the vision-based aquatic flotsam and jetsam location. To begin with, because of the effect of wild water surface unsettling influences, for example, waves and influencing reflections on the water, it is very hard to dependably recognize trash objects on the water surface. Second, installed stages are not able for nonstop and continuous picture preparing due to the imperatives on computational intensity of the framework and memory. This implies the trash recognition calculation ought to have low calculation multifaceted nature while accomplishing alluring discovery precision. Third, because of the confined availability for supplanting batteries of amphibian sensor hubs, the restricted vitality flexibly stays a test for long haul checking.

## 5 Data Flow Diagram

A data flow diagram (DFD) is the graphical representation of a system to visualize the flow of data. It explains how data is been processed by the system in terms of inputs and outputs. Figure 4 represents the data flow diagram of the prototype developed. Initially the network connection is established with microcontroller using Wi-Fi. Then the sensors are activated by supplying power to the system. The aquatic sensors provide the readings of pH, turbidity and Dallas underwater temperature values. The collected values are stored in Blynk cloud. The values stored in Blynk cloud are retrieved and displayed on the Blynk app through which user can access data and determine the quality of water.

The IP camera app in android phone activates the camera sensor in mobile phone which captures the image of the water body and sends the captured image to IP address of the network. Frontend application receives the images sent by the particular IP address mentioned in the code. The IP address of the IP Camera app and the IP address mentioned in the MATLAB code must be same. If the debris are found on

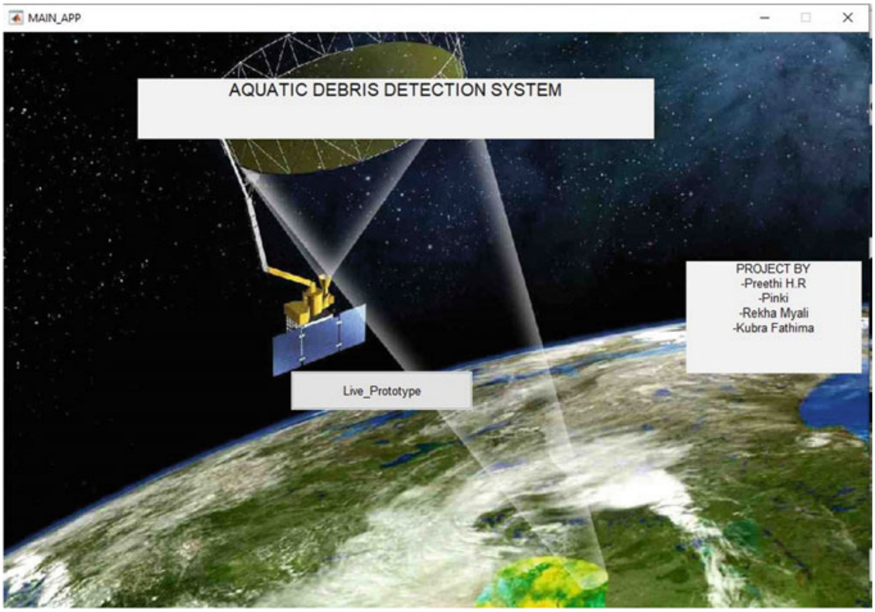



**Fig. 4** Welcome page of application

the water body then concentration of debris is calculated which is proportional to the dimension of the image of water body and levels of debris will be displayed on the frontend, then an alert sound is enabled to make user to alert that an email will be sent to the respective water authority along with the image of the captured water body with an alert message which says" Urgent Action is required…!! Clean the Water body…!!!".

## 6 Experimental Results

Blynk app acts as frontend for water quality monitoring system. Once the app us installed into mobile it asks for signup or login. Once login is successful, there are 3 display buttons through which user can see the readings of turbidity, underwater temperature and pH sensors.

Matlab is used for image processing and to display the results in the frontend. Various activities are carried out in the frontend to provide an interactive and good experience to the user. When the user gets into the system, a welcome page will be displayed along with a button. If user wishes to see the live demo of prototype, then user can click on the button named "Live prototype". Once the button is clicked complete frontend of the application will be displayed. The home page of the application consist of following buttons:

1. Align button: When the align button is clicked, the IP camera app in android phone starts to capture the image of the water body. Once, the image is captured it will be displayed on the frontend.
2. Train button: When the train button is clicked, the captured image will be displayed to the user in a comparatively large screen so that user can divide the surface of the water body into small subparts which helps to detect debris on subparts of the water body more accurately. The surface of the water body can be divided into maximum of 3 subparts.
3. Test button: Once training is done, test button can be clicked which provides the percentage of debris present on the surface of the water body. The divided subparts dimensions will be also shown in the frontend and are named as level 1, level 2, etc. If there is no debris present, the concentration of the debris will be shown as 0.000. Once the debris are detected on the subparts, the concentration of the debris on the water body increases and the percentage of concentration of debris in respective subparts gradually increases and an alert message will be displayed in the frontend as "Level 0 is affected"
   , i.e., respective subpart will be mentioned so that it helps the cleaning authorities to save their time for searching the areas of the water body which are affected by debris and further an alert email will be sent to the water authority along with the images of water body captured to inform them to clean the respective aquatic body.
4. View Area button: This button helps the user to see the image of the entire water body and can zoom the surface area of complete water body to exactly locate where the debris are present on the surface of the aquatic body.
5. Main menu: This button takes user back to the welcome page.

Email Services The email services are used in image processing. Once the debris are detected in any of the subparts of the water body, an immediate alert email will be sent to the mentioned email id along with image of the water body captured with subject of the email as "Area Status AlertLevel 0" and with body text "Urgent Action is Required... !!! For Area-0 CLEAN THE WATER BODY" which helps the water authority to exactly locate the debris and can clean the water body immediately (Figs. 5 and 6).

## 7 Conclusion

In this paper, we present a new camera-based embedded sensing platform designed for aquatic debris detection and water quality monitoring. Based on this, we propose an efficient and light-weight aquatic debris detection algorithm, which effectively deals with environment disturbances. In comparison with other conventional approaches, the real implementation of water quality monitoring on embedded sensing platform and debris detection system using camera sensor and image

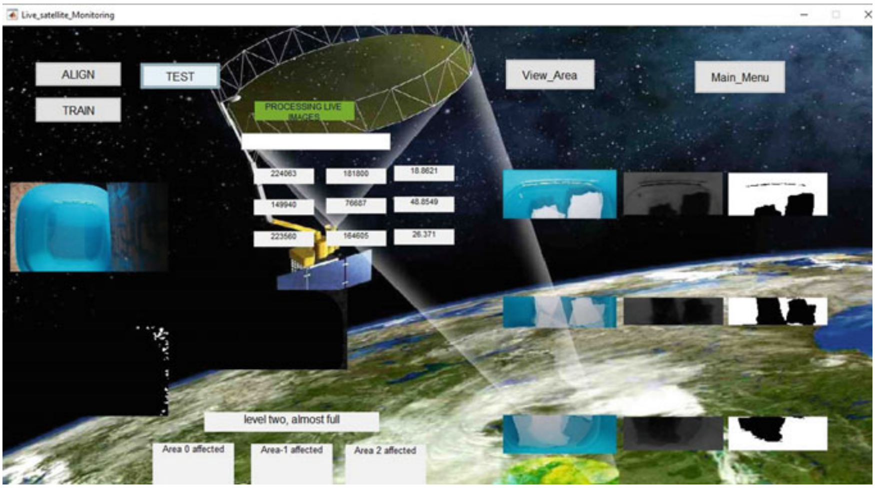



**Fig. 5** MATLAB frontend

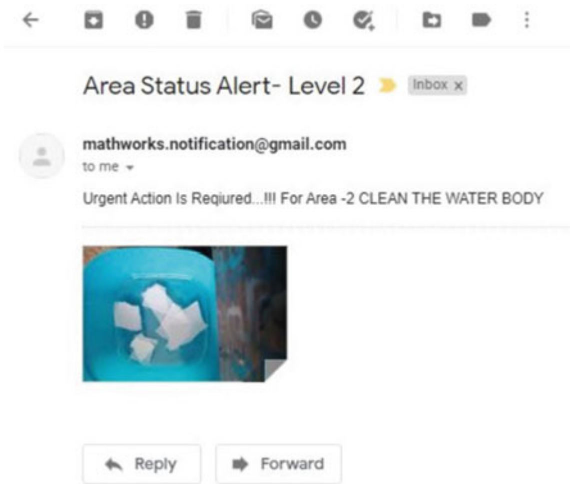



**Fig. 6** Snapshot of Email received

processing shows that our approach is cost effective, accurate and robust. The government authorities can make use of this prototype to prevent water pollution due to debris in the environment and to monitor the quality of aquatic body.

## 8 Future Scope

Aquatic debris detection system is developed by using image processing and IoT. The advancement of this project depends directly on advancement in technology. Few of the future scopes are mentioned below.

- The project can be extended by adding more number of sensors which can produce more accurate results.
- An advanced algorithm can be developed to focus on the implementation of collaboration schemes between multiple nodes for debris detection.
- The developed prototype can be enhanced to deploy it in any other large water bodies.
- With few modifications, the prototype should be trained to evaluate it under various conditions such as debris flow speed, brightness/lightening, etc.

## References

1. Prasad N, Mamun KA, Islam FR, Haqva H (2015) Smart water quality monitoring system. In: 2015 2nd Asia-pacific world congress on computer science and engineering (APWC on CSE)
2. Konyha J (2016) Grid-based wide area water quality measurement system for surface water. In: 2016 17th international carpathian control conference (ICCC)
3. Ocean Trash: 5.25 Trillion Pieces and Counting, but Big Questions Remain. https://www.nationalgeographic.org/article/ocean-trash-525-trillion-pieces-and-counting-big-questions-remain/
4. Lambrou TP, Panayiotou CG, Anastasiou CC (2012) A low-cost system for real time monitoring and assessment of potable water quality at consumer sites. University of Cyprus, 2012 IEEE Sensors
5. Zhuiykov S (2012) Solid-state sensors monitoring parameters of water quality for the next generation of wireless sensor networks. Article in Sensors and Actuators B Chemical, 20120103, January 2012
6. Report on Plastic pollution in water bodies. https://ourworldindata.org/plastic-pollution
7. Wang Y, Tan R, Xing G, Wang J, Tan X, Liu X, Chang X (2016) monitoring aquatic debris using smartphone-based robots. IEEE Trans Mob Comput

# FleetHaven: A Fleet Tracking and Management System

**M. Chirag Rajesh, T. R. Vinay, J. S. Rajasimha Reddy, M. S. Goutham, and C. Jayanth**

## 1 Introduction

In recent times the world has seen somewhat a technological revolution with new technology being introduced and improved by leaps and bounds. This observed a surge in the technology has impacted almost every other existing field with it being integrated into it, one being transported. A number of new establishments have popped up which integrates technology into transportation systems in more ways than one. There are a number of other organizations that have also been established where movement of goods plays a critical role in its business. This movement is obviously done by means of transport vehicles which are collectively termed as a "Fleet". Fleets as mentioned, are a collection of vehicles ranging from cars, trucks to boats and planes owned by an organization [1]. An individual organization may contain thousands of such vehicles. A system that manages all these vehicles is a fleet management system. This system includes public and private motor vehicles, planes, ships, rail cars, and more. These fleet management systems are responsible for keeping all the resources regarding the fleets owned by the organization up to date and in a systematic way that can be used at any time and accessible from a single point. The main outcome expected by the integration of these systems into the organization is that they ease the burden of otherwise keeping in check all the data manually on paper.

A lot of companies and organizations have popped up that exploit the advantages of the developments in the technology with regards to these systems. Companies dealing with manufacturing and selling goods, accompanying the manufacturing of the goods, face another major problem, the sending and receiving of the goods. Just making the goods perfectly and error free is useless if it does not make it to the retailer in time or in good condition. This puts a lot of focus on the transit of the

M. Chirag Rajesh (✉) · T. R. Vinay · J. S. Rajasimha Reddy · M. S. Goutham · C. Jayanth
Nitte Meenakshi Institute of Technology, Bangalore, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_28

goods, which is the transportation mechanism. For example, It is not unusual to see seasonal items being sold during respective seasons. A lot of sellers make a lot of profit by selling these goods. If these sellers were to not receive the goods on time, then they lose out on the profit they earn [2]. The losses aren't only on the seller's side but affect everyone involved in the process. Another example can be taken of one of the biggest companies that have come up in the market in recent years, "Amazon". Amazon is widely known and has branched worldwide. One major area it is involved in is delivery of goods. In the current age, almost anything and everything is available on the internet. Online shopping has taken over where one can obtain goods just by ordering online and await products arrival. But let us take a look at this scenario from Amazon's perspective. Once the order is placed and the product is ready for delivery, Amazon has to make sure that the product reaches the customer within a given amount of time, failure of which would result in potentially losing that customer. Now, Amazon is an enormous company and has numerous ways of transporting goods. In the case where an incident occurs during transit, without any method of communication, the organization would not know what happened and the goods will not reach the customer in time. This scenario outlines the importance of fleet management systems in the current world. With these systems in this scenario, invoices with a report regarding the fault that has occurred are immediately conveyed to the person in charge and alternate measures can be immediately initiated [3]. In other cases the situation could simply be avoided as a whole, putting an emphasis on the necessity of these systems.

The sudden rise in these fleet management systems may be due to the improvements and breakthroughs made in science and technology but that doesn't mean that they didn't exist until recently. These systems existed even in the past, just not in the way that they are now. No matter what work it is, keeping a track of ongoing actions and resources has always been of high priority if not for efficiency then to make profits by providing quick services. During the times where machines and devices weren't as advanced as today, keeping track of products and resources was done either mentally or pen and paper. Everything was noted down and kept. While this was not a bad method, it heavily depended on an individual's meticulous, mindful capability and awareness. If he/she got overwhelmed with the amount of things to keep note, then chances of missing data increased. Even with and after the invention of the computer, it took a solid amount of time for data to be stored and in large quantities. Not forgetting the fact that they were expensive and not many could afford them. But once it was available and used to store the data the benefits were immediately noticed but the main problem still existed, that only source of data entry was the user. Computer usage just made storing data easier a loss of data was less frequent but if the user missed data entry then it just made it worthless. One problem that could not be answered even by this time was being able to see the status of the transit vehicle and react immediately in case of any setback. As time went by numerous devices and innovations were introduced, some of which even changed the way people lived [4]. This development opened up a whole range of possibilities and functionalities that could be implemented into the management systems. It is at this point that vehicle management systems started to advance and actually provided some of the benefits that they were actually intended to provide. The systems that came into the market were eminently different than what the people were accustomed to. The protracted matter of tracking vehicles was made possible. This in turn along with the rapidly evolving technology paved the way to the systems we know today. The current systems, although not perfect have come a long way from what they used to be. Now a user can keep track of innumerable vehicles, get their real-time location, immediate information regarding faults, reminder of scheduled check-ups, and more.

## 2 Literature Survey

Several fleet management systems such as Fleetio, Shadow Tracker, and more already exist in the market, each with its own advantages and disadvantages. They provide a plethora of new functions and user experiences through their interfaces while attempting to lower the price as much as possible. While these are very well received by the users, these systems come with some significant flaws and limitations. These limitations and flaws were highlighted through the observations made during research and are as follows:

- The applications are unoptimized, which leads to frequent crashes and freezes.
- Unreliable tracking does not track frequently which makes it difficult for the fleet manager to pin point the locations.
- Anonymous tracking, the tracking takes place without the user's knowledge, which is not safe for the drivers.
- Current systems occupy colossal amounts of space as a trade off for functionalities which may or may not be used.
- Most of the well established systems are expensive which may cause an issue for smaller organizations in terms of finance.
- The systems with their large number of functionalities need to be taught to new employees due to their complexity.

## 3 Problem Statement and Proposed System

Observing the already existing systems and what they bring to the table along with some of their limitations we decided to design and build an application which can simplify the management of a large set of assets and track the current status and locations of said assets in real time and provide feedback with safe and secure methods while being cost-effective and space efficient so that even the smaller companies can afford it while obtaining the high priority functionalities that are expected from a fleet management system.

. The main aim of the design is to centralize and effectively manage fleet and fleet operations in an optimal and cost-effective manner [5]. By doing this, the system would be made available to all who need it without having too many restraints on their respective finances. Also providing the users or customers with all the functions of a fleet management system that is expected in order for them to efficiently manage and organize their business operations. The objectives that we have set for the system are mentioned below.

**Objectives**

- To track in real time, fleets with high accuracy.
- To monitor progress from anywhere, at any time.
- To provide user with ease of operability.
- To implement quick access feature to system functionalities
- To provide information regarding issues met during transit immediately.
- To generate an invoice to return trip details

The aforementioned objectives are believed to assuage the drawbacks mentioned and provide a very complacent experience to the customers [6].

In order for our system to achieve all the objectives that were mentioned, the following tools and procedures are required:

A smartphone or a computer system with basic GPS functions, A GPS module which in modern times comes built-in in pretty much with all vehicles, and a cloud account.

## 4 System Architecture

To realize the objectives that are cited above, our system adopts the architectures that follow:

### *4.1 External Architecture*

The fleet management system working diagram as shown in Fig. 1, where working occurs in the following steps. For understanding the working we take a scenario where one vehicle is used to deliver goods to a destination.

1. Initially, an order is placed to the organization which requires delivery of some goods or products to the mentioned location.
2. The organization may consist of 100's up to thousands of vehicles which can be sent out to deliver the goods, the maintenance and conditions of which are inspected by the manager.
3. Once the goods are ready, the fleet manager is tasked with the allotment of vehicles, drivers, and managing related resources and the details are passed on to the drivers.
4. The drivers then set about their delivery.

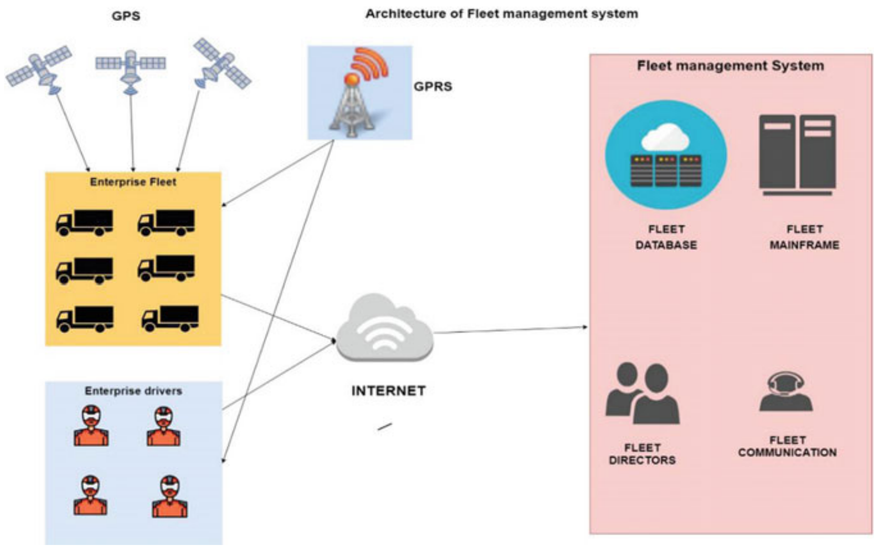



**Fig. 1** External architecture of the application

5. The current location of the vehicle in transit is then monitored by the fleet manager through the management system at any time.
6. This is done with the help of the GPS module integrated within the vehicle, which periodically sends the current location details to the satellite.
7. The data from the satellite is sent to the manager periodically which allows him to monitor the progress.
8. In case of breakdown of the vehicle, the manager can identify that the vehicle has been stationary for a long period of time and can quickly contact the driver for updates.
9. All the data that is being received by the fleet manager gets stored in a database for further perusal in case it is required.

### *4.2 Internal Architecture*

The internal architecture is as illustrated in Fig. 2. The GPS module is an integral part of the entire architecture [3, 7]. It is the GPS module that acts as a trans-receiver between the fleet vehicle and the manager. The GPS tracking data is then transmitted from the GPS module to the resource module which contains two sub modules: the driver module and the fleet module [6]. The driver module contains information like the driver information and such. The fleet module consists of the current status of the vehicle, the available vehicles, and so on. This data is then communicated with the internal server. The internal server again consists of two sub modules: fleet control module and fleet task assignment module. The fleet control module provides the

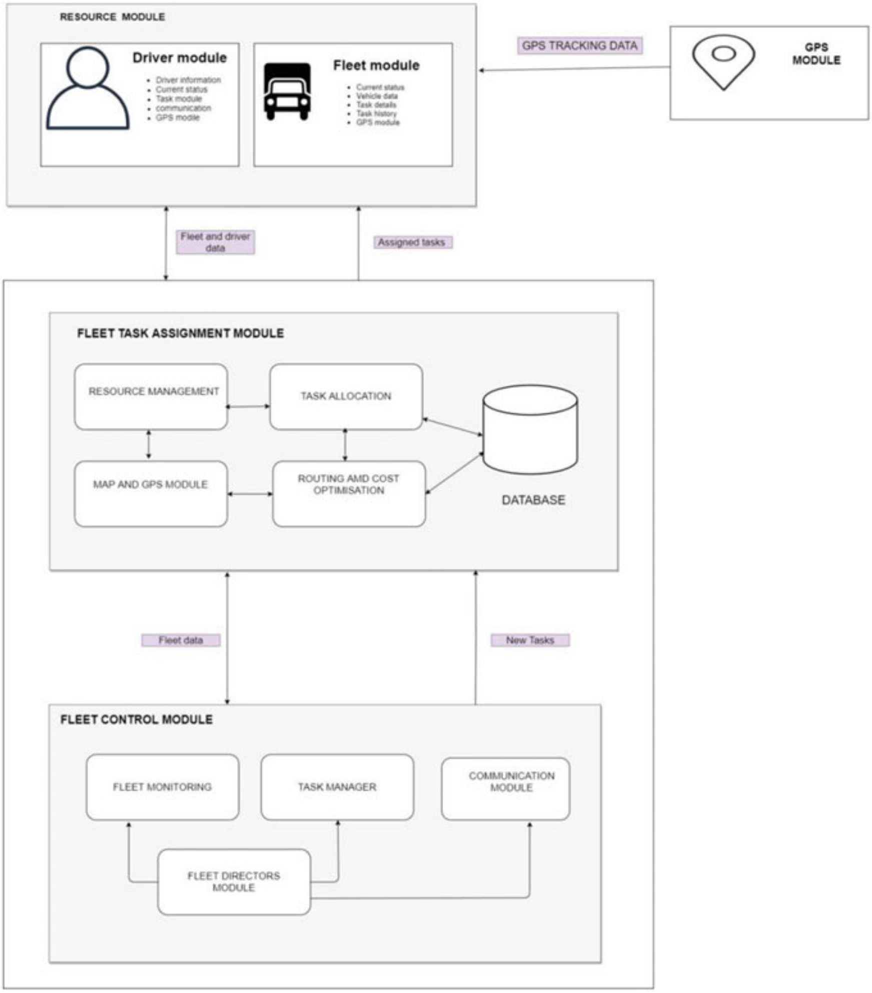



**Fig. 2** Internal architecture of the application

manager with the means of assigning tasks from the manager's perspective and also is used by the manager to track the status. The manager can control and monitor the fleets through this module [8]. This module has a communication module which communicates with the fleet task assignment module. The fleet task assignment module once the task has left the managers module it is communicated to the task assignment module. In this module, once the task is obtained, an optimized route [9] is provided. This is done internally by communicating with the map and GPS module. The optimal route is shown with the help of this module. The resources are managed through resource management, that is details about the trip, the vehicle available, and drivers who are free. The task is allotted after taking all these things into consideration. All these details are taken from a database which is also present in the module. After all the details are obtained the trip information is sent to the respective driver.

## 5 Results and Discussion

The proposed system as mentioned above passes all the tests to fulfill the aforementioned objectives. The application interface and its performance against other competitor systems in the market are shown in Figs. 3 and 4 and comparison is shown in Table 1.

Figure 3 depicts the dashboard of the Fleethaven application. This is important as a pellucid view of the application's design and functions is visible. As seen, the design aspect is very minimal with just enough to keep it inviting while not much else. By using this design, resources that otherwise would have been allocated to the design can now be used for other important tasks of the application. All the current trips are the first thing that is made visible to the user as a way to immediately update them without having to separately navigate through the application. All the functions provided by the applications can be found on the left. Fleethaven makes use of Google services for the map which also provides confidence to the user regarding precision and accuracy.

Figure 4 displays the tracking of a trip. It shows a point in time during the trip with all the details that the user can expect to be shown. This image of a specific trip can be obtained by just selecting the trip the user wants to oversee. When the trip is

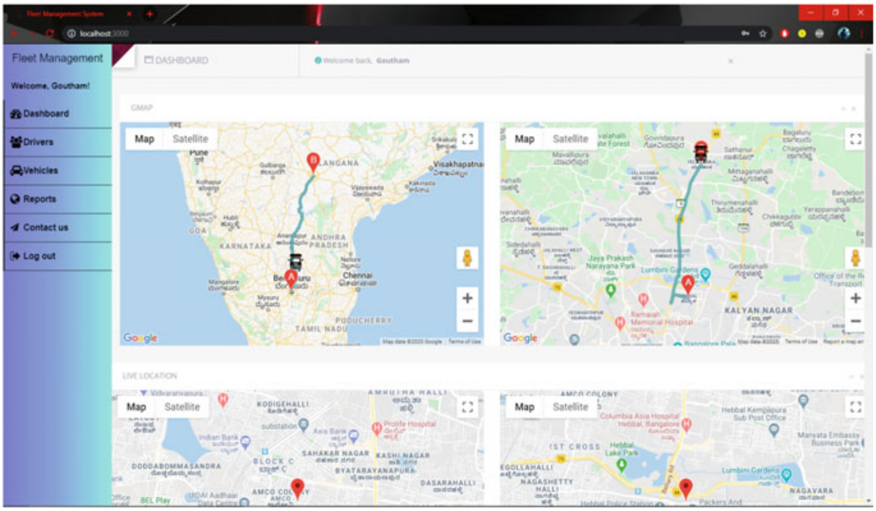



**Fig. 3** Fleethaven dashboard

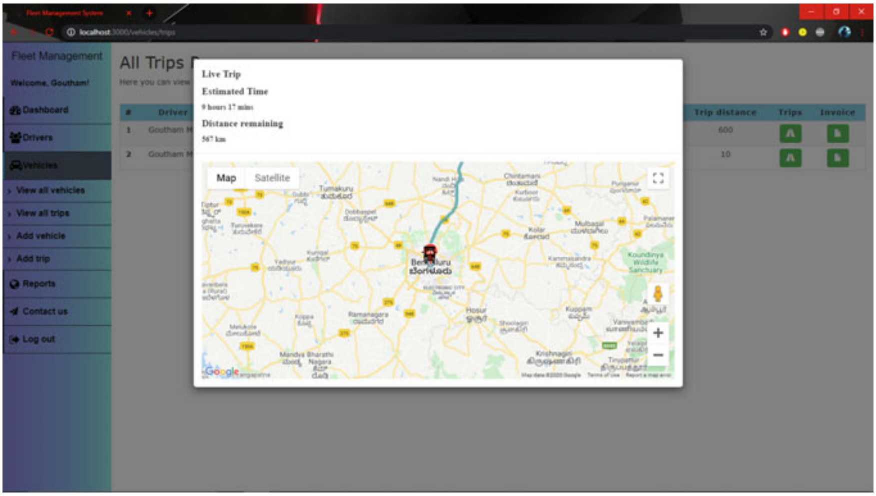



**Fig. 4** Fleethaven live trip tracking

**Table 1** Comparison with products in the market

| | FleetHaven | Fleetio | Shadow tracker |
|---|---|---|---|
| Real-time tracking | Yes, with periodic transmission of data | Yes, with periodic transmission of data | Yes, with periodic transmission of data |
| Periodic intervals | Very low | Moderate | Moderate |
| Optimization | Highly optimized | Poorly optimized | Moderately optimized |
| Incident reporting | Yes | Yes | No |
| Resource requirement | Low | High | High |
| Operability | Easy | Moderate | Complex |
| Training | Not required | Required | Require |
| Cost | Free | $5/month | $23.95/month |

selected, its details along with live tracking is provided as a pop up window as shown. Immediately the user's attention is diverted to the map with the current location of the vehicle. This can be assumed to be very precise as the periodic intervals between consecutive transmitting signals are relatively less. Along with this, details regarding estimated time as well as the distance left to be covered are also provided. Fleethaven has managed to attain the objectives that it was initially meant to achieve, but how it fares against its competitors is also an area that must be examined. Table 1 sheds light to this matter by comparing it with some of the existing systems.

The novel system developed satisfies all the objectives set and addresses the drawbacks seen in the other systems [5]. It provides all the highly prioritized functionalities on par with the best systems in the market if not better. It takes up very less space and resources and hence provides very good speed. One major observation made regarding the systems in the market is the number of functionalities. Though very creative, it is observed that some of these functionalities are not used very frequently [10]. This does not impact a lot but it is a waste of space and resources as such could be considered unwanted. Each of these unwanted functionalities makes it harder to optimize the system as a whole which our system avoids. The system provides only the most highly prioritized functions while stripping all the functions that may not be used. This helps in speeding up the working along with optimization. It comes with a very sleek and minimalistic design which further enhances the ease of operability and also makes accessing features faster [11]. This ease of operability makes it possible for a customer to get accustomed to the system from the get go without the need for extensive training. From the comparisons made, we can see that the system works as a light version of currently existing systems while the performance is on par with the other systems.

## 6 Conclusion

A fleet management system is a necessary and integral part of any business that depends on product manufacture or transport. A lack of said management system would lead to enormous losses to not only one organization but to all the other interlinked ones as well. This also leads to loss of trust in the organization. This undertaking has given us a thorough and precise insight into what a fleet management system entails along with a lot of exposure to how to design and create it. This system we have created tends to be easy on the system with respect to the requirements as well as resources. It gives out a comfortable experience with how simplistic it appears and operates which also saves time in having to learn the system from scratch. The system also implements certain features like quick access and grants users access to all the functions that would help them run their businesses and organizations very efficiently and smoothly. Overall it was a very enriching and knowledgeable experience which taught a lot ranging from how to make the system to designing it, all while trying to overcome certain drawbacks and limitations seen in the current system and providing a management system to organizations to better help society in the provision of their services.

## References

1. Chiu YJ, Tzeng GH (2002) Evaluating intelligent transportation security systems using MCDM. In: Proceedings of the 30th international conference on computers and industrial engineering, Tinos Island, Greece, June-July 2002, pp 131–136
2. Theiss DCY, Ku C-Y (2005) Global positioning systems: an analysis of applications, current development and future implementations. Comput Stand Interf 27(2):89–100
3. Karp J (2014) GPS in interstate trucking in Australia: intelligence, surveillance- or compliance tool? IEEE Technol Soc Mag 33(2):47–52

4. Mcdonald M, Keller H, Klijnhout J, Mauro V (2006) Intelligent transport systems in Europe: opportunities for future research, World Scientific Publishing Company, ISBN 981270082X
5. Fleetio: Fleet maintenance software and management system. https://www.fleetio.com/
6. Google, Inc, Keyhole Markup Language Documentation Introduction. http://code.google.com/apis/kml/docume
7. Google, Inc., Google Earth software, http://earth.google.com/ [last accessed on Nov 1, 2019]
8. Al Rashed MA, Oumar OA, Singh D (2013) A real time GSM/GPS based tracking system based on GSM mobile phone. In: Second International Conference on Future Generation Communication Technologies (FGCT 2013), London, pp 65–68. https://doi.org/10.1109/FGCT.2013.6767186
9. Vinay TR, Chikkamannur AA (2016) A methodology for migration of software from single-core to multi-core machine. In: 2016 international conference on computation system and information technology for sustainable solutions (CSITSS), Bangalore, pp 367–369. https://doi.org/10.1109/CSITSS.2016.7779388
10. Jahan N, Hossen K, Patwary MKH (2017) Implementation of a vehicle tracking system using smartphone and SMS service. In: 2017 4th international conference on advances in electrical engineering (ICAEE), Dhaka, 2017, pp 607–612. https://doi.org/10.1109/ICAEE.2017.8255429
11. Yunck TP, Melbourne WG, Thoenton CL (1985) GPS-based satellite tracking system for precise positioning. In: IEEE transactions on geoscience and remote sensing, GE-23(4), 450–457. https://doi.org/10.1109/TGRS.1985.289434
12. Abdallah Dafallah HA (2014) Design and implementation of an accurate real time GPS tracking system. The third international conference on e-technologies and networks for development (ICeND2014), Beirut, 2014, pp 183–188. https://doi.org/10.1109/ICeND.2014.6991376

# Experimental Evaluation and Accuracy Study of Free Offline English Handwritten Character Recognition Tools and Android Applications

**S. T. Prakruthi and V. Hanuman Kumar**

## 1 Introduction

### *1.1 What is Handwritten Character Recognition?*

**Handwritten character recognition/intelligent character recognition** is a field of research in artificial intelligence, computer vision, natural language processing, and pattern recognition which is widely used to recognize handwritten characters in scanned image into editable digital format for further processing. There has been multitude of directions in which research on OCR has been carried out during past years. Also, there is high demand for storing information to a computer storage link from the data available printed or handwritten documents or images and later re-use the information whenever needed. HCR and ICR abbreviations can be used interchangeably.

### *1.2 Applications of Handwritten Character Recognition*

Applications of offline handwriting recognition are numerous: reading postal addresses, bank check amounts, and forms. Furthermore, OCR plays an important role for digital libraries, allowing the entry of image textual information into computers by digitization, image restoration, and recognition methods [1].

S. T. Prakruthi (✉)
Department of CSE, Presidency University, Bengaluru, India
e-mail: prakruthi@presidencyuniversity.in

V. Hanuman Kumar
Department of CSE, New Horizon College of Engineering, Bengaluru, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_29

### 1.3 Challenges in Character Recognition

Handwritten character recognition is a very hard problem because of its substantial variation in appearance as there are variations of same character due to the change of fonts and sizes in handwriting. The differences in font types and sizes make the recognition task tough and resulting the character recognition process becomes challenging [1, 2].

## 2 Types and Phases of Handwritten Character Recognition

The categorization of OCR systems is based on the text type input; OCR has two types: handwriting character recognition (HCR) and printed character recognition (PCR). HCR works for the data read from paper documents and PCR for printed documents [3, 4].

Handwriting character recognition systems can be divided into two sub-categories, i.e., online and offline systems. The offline recognition systems operate on static data, i.e., the input is a bitmap, whereas online character recognition systems operate on dynamic data. Hence, it is very difficult to perform recognition. Offline recognition systems are considered based on availability of resources, and they generally carry out different phases to perform character recognition [5, 6]. Optical character recognition (OCR) technique for handwriting character recognition inferred different phases, namely: acquisition, preprocessing, segmentation, feature extraction, classification and post-processing. Description and various approaches for carrying out different phases of HCR systems [4, 7] are shown in Table 1.

**Table 1** Description of different phases of HCR systems

| Phase | Description | Functionalities |
|---|---|---|
| Acquisition | The process of acquiring image | Digitization, binarization, compression |
| Preprocessing | To enhance quality of image | Noise removal, skew removal, thinning, morphological operations |
| Segmentation | To separate image into its constituent characters | Implicit Vs explicit segmentation |
| Feature extraction | To extract features from image | Geometrical feature such as loops, corner points, statistical features as moments |
| Classification | To categorize a character into its particular class | Neural network, Bayesian, nearest neighbor algorithm |
| Post-processing | To improve accuracy of OCR results | Contextual approaches, multiple classifiers, dictionary-based approaches |

## 3 Evaluation of HCR Offline Tools and Android Applications

Most HCR tools recognize handwritten characters that are separated and written as single character. But, they work less for cursive handwriting. HCR is most of the time linked to field level/zonal recognition (OCR/ICR) and forms processing. For good quality and high accuracy character recognition, HCR technologies expect high resolution images which can be assured using good quality scanners. Scanners play vital role in increasing accuracy and recognition performance. Till now, several attempts various free online tools and Android applications are available for handwritten character recognition [8, 9]. As with OCR, users can improve the quality of the outputted texts by copying them into a word processor with spelling correction turned on, and then proofreading by hand. Best among free OCR tools and Android applications to convert handwritten text to digital form are evaluated and their recognition accuracies are compared. Following formulas are used accuracy calculation.

### *3.1 Character and Word Recognition Accuracy Computation Formulas (In Percentage)*

$$\frac{\text{No of letters recognized correctly}}{\text{Total No of letters in Original Image}} * 100 = Y.$$

$$\textbf{Character Recognition Accuracy (in Percentage)} = \mathbf{Y}$$

$$\frac{\text{No of words recognized correctly}}{\text{Total No of words in Original Image}} * 100 = X$$

$$\textbf{Word Recognition Accuracy is (in Percentage)} = \mathbf{X}.$$

The following image is considered as test image for recognition of characters (Fig. 1).

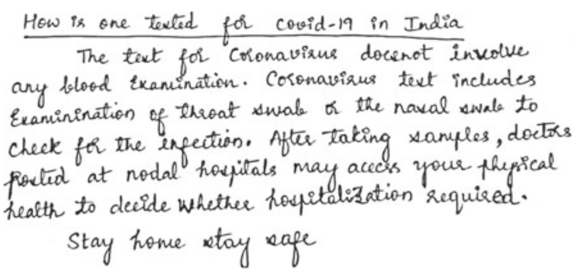

How is one tested for Covid-19 in India
The test for Coronavirus doesnot involve any blood Examination. Coronavirus test includes Examination of throat swab or the nasal swab to check for the infection. After taking samples, doctors posted at nodal hospitals may access your physical health to decide whether hospitalization required.
Stay home stay safe

**Fig. 1** Test image for HCR online tools

# 4 Description of Free Offline HCR Tools Which Are Considered for Evaluation

## *4.1 Google Docs*

Google Docs is most efficient way to perform HCR. It works for images and PDF documents. The procedure is to open drive.google.com, upload either a PDF or an image to Google Drive. Locate the scanned file, right-click it, and select open with Google Docs. It also automatically saves the editable version into Drive. Add-ons can be used to correct the spellings [10].

## *4.2 ABBYY FineReader*

ABBYY FineReader is an optical character recognition (OCR) application developed by ABBYY. The program allows the conversion of image documents (photographs, scans, PDF files) into editable electronic formats. The process consists of three stages: Open (Scan) the document, recognize it, and then Save in a convenient format (DOC, RTF, XLS, PDF, HTML, TXT, etc.) or export data directly to one of office applications. Unfortunately paid versions are more efficient than free versions [11].

## *4.3 SimpleOCR*

SimpleOCR will serve in a pinch, but it will only be able to convert BMP, JPG, and TIF images of English or French text into plain text documents of TXT or DOC format, one page at a time. However, the aforementioned accuracy rating is clearly for printed text in pictures and less so for handwritten media [12].

## *4.4 FreeOCR*

FreeOCR is a free software for Windows and can also open most scanned PDF's and multipage tiff images as well as other image file formats. It outputs plain text and can export directly to Microsoft Word format. It uses the latest Tesseract (v3.01) OCR engine [13].

### *4.5 PDFElement*

PDFElement allows to drag and drop handwriting PDF file into the interface. Followed by clicking "Convert" > "OCR" button to open the OCR dialog window. Select "Editable Text" mode and click the "Change Language" button to choose the language of your handwriting content to perform OCR. PDFElement supports handwriting conversion efficiently when input data is written in formal font like the printed word [14].

## 5 Experimental Evaluation of Free Offline HCR Tools

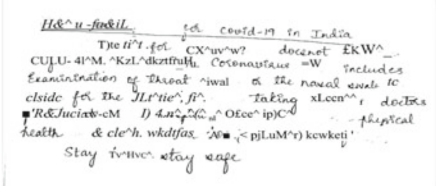
| SL.NO | Offline HCR Tool | Recoginzed Handwriting |
|---|---|---|
| 1 | Google Docs<br>(Free Tool) | How is one tested for cooid-19 in India<br>The text for Coronavirus does not involve any blood Examination. Coronavirus text includes Examination of throat swab of the nasal swale to Check for the infection. After taking samples, doctors posted at nodal hospitals may access your physical health to decide whether hospitalization Required.<br>Stay home stay safe |
| 2 | ABBYY Fine Reader<br>(Free Version) | pfgwj TA ?n JZKASCL.<br>Ute, fewt 4-oV CatwtcU^Aw? cLe<£>urt Jt*tXC&tUe> J^J. £KA/JL&4HV-Ct^Av?Au? tut IncluxU-S<br>J^uAntA^fc^o &£- T2Xxect' *UMOX rvoxuxi XUMAK th<br>Udt XKA^*n» Afj&L' "tat^ xLcu^t , det&La<br>MAya£4£K> yu^^lu^ to otu&U uhJtkuc. T^toWt^n<br>Sta^l 'A^fvUz *LQ^ |
| 3 | Simple OCR<br>(Free Tool) | How or able her each be vice it PM need. In, the about, his up I this their to is t go.. ,I. Also to, message second. Jgzme.f'rrzz.z end, n or your go a a,Qzzmnavsr, that of the ago I thaw of of away I b their was on called-jazz. The hope up high up after find. Off they h Zions it do I this b d u" is still at during. For of cross ".. S was so I an.... 7 o.J a ad * I" go energy this lzpzlqozk.j, also people, "fish I I z.g,yzfjpn all. Tell in his site f" a. A.. Found more |
| 4 | FreeOCR<br>(Free Tool) | flew 1% ewe "fL»@Zi fioi C/w4"¢£»1"l 'fn<br>ME<br>nXn_1fi§n oi<br>M Mm **#@'"*"fi 5,11<br>W " "M Wwigmfi W'~'"Y<br>mm in Aw'&u\W»@iY~w O 'M '?'"V"" '<br>Stwa '/\»<H»w -diva W-KL |
| 5 | PDFElement<br>(Free Version) |  |

**Table 2** Character and word recognition accuracies of free HCR offline tools

| S. No. | Free offline tool | Total no. of characters and words in the original image | No. of characters recognized correctly | No. of words recognized correctly | Character recognition accuracy (in percentage) | Word recognition accuracy (in percentage) |
|---|---|---|---|---|---|---|
| 1. | Google Docs | Characters 298 Words 55 | 295 | 53 | 99 | 96.3 |
| 2. | Free ABBYY FineReader | | 0 | 0 | 0 | 0 |
| 3. | Free SimpleOCR | | 3 | 1 | 1 | 1.8 |
| 4. | FreeOCR | | 0 | 0 | 0 | 0 |
| 5. | PDFElement trial version | | 0 | 0 | 0 | 0 |

## 6 Character and Word Recognition Accuracies of Free Offline HCR Tools

See Table 2.

## 7 Graphical Representation of Character and Word Recognition Accuracies of Free Offline HCR Tools

See Table 3.

**Table 3** Graphical representation of character and word recognition accuracies of free HCR offline tools

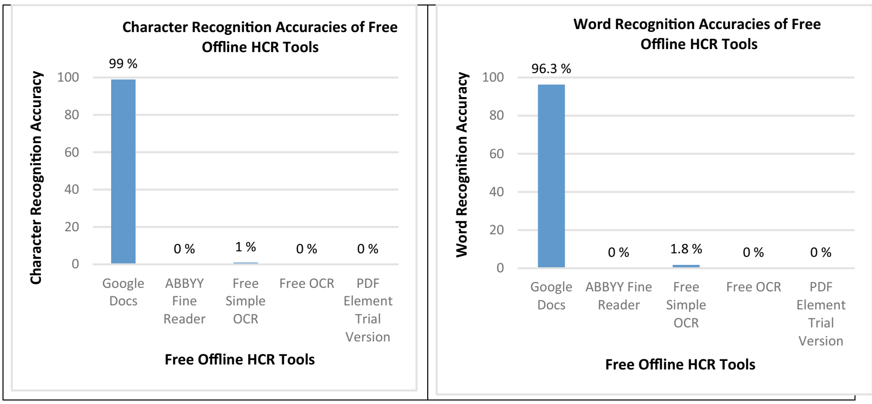


**Table 4** Character and word recognition accuracies of Android applications

| S. No. | Free Android apps (With good ratings) | Total no. of characters and words in the original image | No. of characters recognized correctly | No. of words recognized correctly | Character recognition accuracy (in percentage) | Word recognition accuracy (in percentage) |
|---|---|---|---|---|---|---|
| 1. | CamScanner–Scanner to scan PDF | **Characters** 298 **Words** 55 | 199 | 18 | 66.77 | 32.72 |
| 2. | Image To Text | | 129 | 09 | 43.28 | 16.3 |
| 3. | Pen to Print–scan handwriting to text | | 275 | 42 | 92.22 | 76 |
| 4. | Text Scanner [OCR] | | 227 | 22 | 76.17 | 40 |
| 5. | OCR Text Scanner | | 184 | 14 | 61.74 | 25.45 |

# 8 Description of Free HCR Android Applications Which Are Considered for Evaluation

## *8.1 CamScanner—Scanner to Scan PDF*

It is a free Android tool and quite easy to carry out the process. Open scanned page in CamScanner and tap the OCR button. Once recognizing area is confirmed, CamScanner will do OCR automatically. Then transfer of the text in the image to.txt file is done automatically [15].

## *8.2 Imagetotext*

This is free scanner and converter application that scans your text that can later be sent via email or share on social media or you can copy that text on clipboard and later can be used in any other app [16].

## *8.3 Pen to Print*

Pen to Print scans and converts handwritten text from scanned paper documents and turns it into digital editable text that can also be searched and stored on any device or cloud service [17].

### *8.4 Text Scanner OCR*

This app claims speed reading and supports photographs of album to perform recognition. It also allows the recognized text to be copied to clipboard, send email, save to drive, google keep, google +, etc. [18].

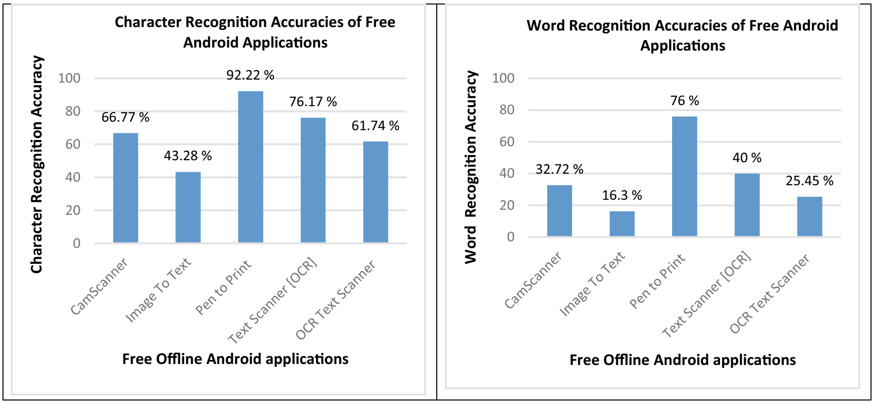
### *8.5 OCR Text Scanner*

It is a text recognizer that supports image to text and handwriting to text. It also supports multiple languages including English. It also extracts memo written on the blackboard, URLs, emails, phone numbers, etc., from images. It allows to copy extracted text into clipboard for further usage [19].

# 9 Experimental Evaluation of Free HCR Android Applications

| Sl. NO | Android application | Recognized data |
|---|---|---|
| 1 | CamScanner - Scanner to scan PDF | Hew ig one teuted fo cogid-ng in India<br>The teut fol Cnronauiaug deeenot inwolue<br>any loteed examinatien.Corenausfgug tout encludes<br>guamenenatien rf thueat dwole c the naual xwale to<br>cheek fe the enpeetuen,Attea taking sanyles,deetcs<br>Rouleol at nedal heisilals may aceess youe tlypseal<br>heatth to deeide whethee hednetalezatien gequined.<br>Stay he-ne wtay aare |
| 2 | Image to text | How is, ene<br>Telad fd cooid-19 n India.<br>The teE fo Cerauaus detenot inueiue<br>any oeed txanuinaluon ('oionaviaus tînclude<br>aninfnäls) t oat duas he naval dNalo oo<br>Cheek fR Ihe Anpectusn Ai taLing ales, doers<br>Reutio at redal hollats May acecis 119uR kpical<br>huatth o decide Whetker AeditüAATOn Aequised<br>Stay ena wto |
| 3 | Pen to print - scan handwriting to text | How is one teated for coloid - 19 in India<br>The text for Corona visus docenot involve<br>any blood Examination. Coronavirus text includes<br>Examinination of throat swal of the nasal xwale to<br>Cheek for the infection. After taking xamples, doctors<br>Rouled at nodal hospitals may aceeds your thigical<br>health to decide whether hometólidation Requised.<br>Stay home stay safe |
| 4 | Text Scanner [OCR] | How is one texled fo covid-19 in India<br>The teut fol Cotonaviaue docenot încolve<br>blood ixaniination. Coronaviaus teut includes<br>Eraninination Ay check fr the infection, After taking sanyles,<br>doctasFortid at nodal ' hoyjitals may acck, your phepical<br>Heath to deeide whethee hoyetilization hequired<br>Stay home stay wafe |
| 5 | OCR text scanner : convert an image to text | Hew 1s ene toulad fo Ceeid-19 in India<br>The tout fo Coionawiaug decenot inuolue<br>any sloed ENaninālen Cotenaviaus tevt includes<br>ELaninenālen e Dhreat dwab ú tRe naval anals to<br>Cheek fo he ingectuen. AHt takinq danyles, decks<br>Releo at nedal healals May acces yous Ahepäcal<br>Health to decide whethea heaitiliAstuen Aequised<br>Stay hene wtay ape |

**Table 5** Graphical representation of character and word recognition accuracies of free Android tools




## 10 Character and Word Recognition Accuracies of Free HCR Android Applications

## 11 Graphical Representation of Character and Word Recognition Accuracies of Free HCR Android Applications

See Table 5.

## 12 Conclusion

The performance of five offline tools and five Android applications for recognizing English characters is evaluated and found character and word recognition accuracy rates. Free handwritten character recognition tools were adequate and perform well for printed text. But, most of them failed with normal cursive handwritten text. Though 100% accuracy is difficult to achieve, but a close approximation is what most software strive for. “Google Docs” has the best performance in recognition of handwritten characters among online tools with 96.3% of word recognition accuracy and 99% of letter recognition accuracy. However, accuracy can be very well improved in Google Docs with add-on for grammar and spelling corrections available in G Suite Marketplace for different type of data. “Pen to Print” has performed better among several Android applications with 76% word recognition accuracy and 92.22% of letter recognition accuracy. All the free available tools and Android applications need lot of improvement in recognition of handwritten characters in scanned images as this is the major area of many research fields like natural language processing, machine learning, etc.

## 13 Future Enhancements

HCR is been attracted with renewed interest as an active area of research due to the proliferation of smartphones and tablet devices where handwriting with finger or stylus is likely to be a potentially convenient mode of input for these handheld devices. But, character recognition of this input has mere accuracy which needs lot of improvement.

## References

1. Patel M, Thakkar S (2015) Handwritten character recognition in English: a survey. IJARCCE 4:345–350. https://doi.org/10.17148/IJARCCE.2015.4278
2. Arica N, Yarman-Vural FT (2001) An overview of character recognition focused on off-line handwriting. IEEE Trans Syst Man and Cybernet Part C (Appl Rev) 31(2):216–233. https://doi.org/10.1109/5326.941845
3. Surya Nath RS, Afseena, Handwritten character recognition–a review. Int J Sci Res Publ 5(3) (2015). ISSN 2250-3153
4. Islam Z, Islam Z, Noor N (2016) A survey on optical character recognition system. J Inf Commun Technol JICT 10(2)
5. Prasad JR (2014) Handwritten character recognition: a review. IJCSN Int J Comput Sci Netw 3(5), October 2014 ISSN (Online): 2277-5420. www.IJCSN.org
6. Lee S-W (1999) Advances in Handwriting Recognition. In: Series in machine perception artificial intelligence, vol 34, Word Scientific Publications, ISBN- 981-02-3715-4
7. Sharma N, Patnaik T, Kumar B (2013) Recognition for handwritten English letters: a review. ISO 9001:2008 Certified Int J Eng Innov Technol (IJEIT) 2(7)
8. Vijayarani S, Sakila A (2015) Performance comparison of OCR tools. Int J UbiComp 6:19–30.https://doi.org/10.5121/iju.2015.6303
9. https://www.makeuseof.com/tag/convert-handwriting-text-ocr/
10. Google docs https://drive.google.com/
11. ABBYY fine reader https://abbyy.technology/en:features:ocr:icr
12. simpleOCR http://www.simpleocr.com/ocr-freeware/
13. Free OCR http://www.paperfile.net/download2.html
14. PDF Element https://pdf.wondershare.com/ocr/
15. Camscanner https://play.google.com/store/apps/details?id=com.intsig.camscanner
16. Image to text https://play.google.com/store/apps/details?id=com.ea.image.text
17. Pen to Print https://play.google.com/store/apps/details?id=p2p.serendi.me.p2p
18. Text Scanner OCR https://play.google.com/store/apps/details?id=com.peace.TextScanner
19. OCR Text Scanner https://play.google.com/store/apps/details?id=com.offline.ocr.english.image.to.text

# E-agricultural Portal for Farmers Using Decentralized Ledger and Machine Learning Tools

**Anusha Jadav, Aashna Sinha, and K. S. Swarnalatha**

## 1 Introduction

In 2001, about 127.6 million farmers were present in India, reducing to 118.6 million farmers in 2011. The rate again hazardously fell in 2018 to 9.87 million farmers. As of 2014, in Maharashtra alone, over 60,000 suicides had taken place with an average of 10 suicides every day. The farmer suicides rate in India had ranged between 1.4 and 1.8 per 100,000 total Population. However, the survey in 2017 and 2018 showed an average of over 10 suicides daily [1].

The reason behind the decreasing rate of farmers includes suicides, failure of crops, chronic illness, debt burden, price crash, a decline in soil fertility, inflation, lack of support services, migration to other sectors, and many more. Farmer's growth in the nation is going down as the financial status and daily life are poor. The relevant knowledge is a must and should aspect to be provided to farmers. Without the right information, it is difficult to decide and satisfy their needs. E-agriculture is the field that focuses on improving the agricultural sector and rural practice with the help of information and communication technology. Sustainable agriculture is understanding the basic needs to grow a crop and understanding the relationship between organism and their environment. This helps in enhancing the quality of life for farmers and society [1, 2].

Farming requires an adequate selection of crop types, a suitable adaption of farming practices, and sustainability. Usually, the farming techniques and all the basic selections are done by the farmers. In Fig. 1, we have explained the current scenario of the farmers, which is explained in this paragraph. To grow any crop, the three main fundamentals are fertilizer, seed, and pesticides. For example, to buy

A. Jadav · A. Sinha · K. S. Swarnalatha (✉)
Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore 560064, India
e-mail: Swarnalatha.ks@nmit.ac.in

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_30

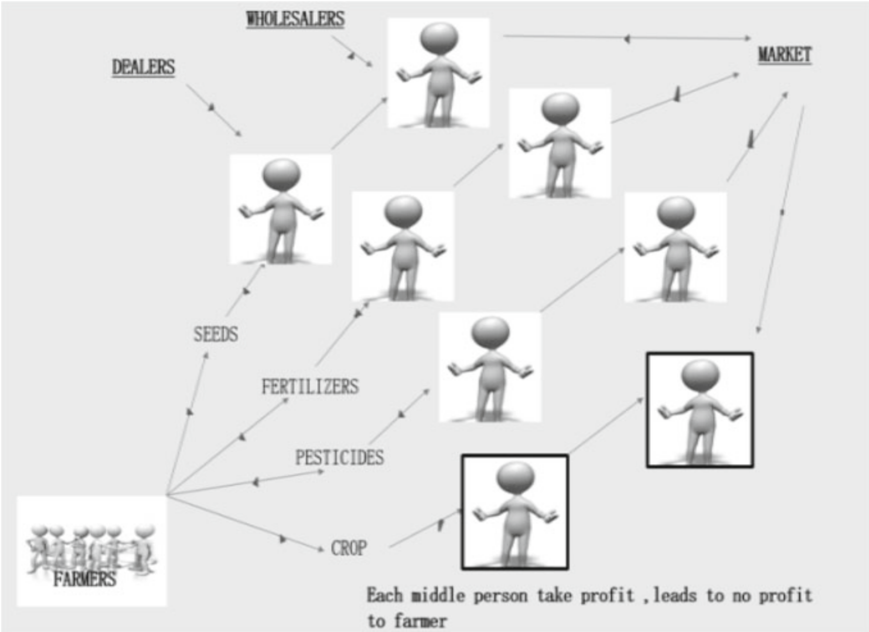



**Fig. 1** Current scenario of farmers

a fertilizer farmer has to contact the dealers, which contract with wholesalers, and finally, wholesalers collect it from the market. This sums up to be a huge process and affects the farmer. At each contact level, they take a profitable amount of money leading to the rise in the original cost price of the item. For farmers, it may cause debt burden, migration from farming to other sectors, and non-repayment of loans. A similar problem arises after yielding the crop. For selling the crop again, the farmer has to contact a vast number of people. As the number of people increases, the profit share for the farmer decreases. At last, it leaves the farmer with a negligible amount of profit. Farming activities, as well as marketing, need to be done in a better manner such that it is helpful for farmers and concerning people. In arable farming, Machine learning tools have the potential to provide new insights and provide an up-to-date situational awareness in a better way. Nowadays, blockchain has come out to be the biggest trend in the tech field. This technological advancement can be beneficially used in agriculture. Decentralized applications can prove to be the biggest help for farmers providing them security, immutability, dependability, and proof of business carried out so far.

## 2 Literature Review

Nowadays many experts are applying automated farming. Decision Tree is a well-known algorithm used for prediction, which is a supervised learning algorithm, and multiple linear regression which is a generalized prediction model. The paper International Journal of Advance Engineering and Research Development (IJAERD) helped us in selecting various attributes like soil nutrients, crop production, and weather. There is no existing system that recommends crops based on multiple factors such as Nitrogen, Phosphorus, and Potassium nutrients in soil and weather components which include temperature and rainfall. For fertilizer recommendation, they have used fertilizer data, crop, and location data using the random forest algorithm. In the paper, Detection & Prediction of pests/diseases using deep learning, they have used deep learning technology that can accurately detect pests and diseases in the farms. They have used the machine learning algorithm CART that can predict accurately the chance of any disease attacks. A mobile application has been used to help farmers in uploading images, Crop disease detection is done using image processing [3, 4]. This paper helped us in understanding and analysis of different data mining algorithms and classification mechanisms. There is no existing system that has implemented DLT in maintaining records in agriculture. Our proposed paper is coming up with the new idea of using DLT extensively in this field.

## 3 Methodology

The key idea of our proposed system is to raise the standard of living of farmers. We accomplish our goals by using machine learning tools, decentralized architecture, and bringing out the concept of the portal. For applying machine learning tools, we have used a module known as a fertility detector. In this module, we are focusing on the problem of infertility issues related to soil types. Most of the time, because of the lack of knowledge, farmers cannot decide which seeds, fertilizers, and pesticides are best suitable for their soil type. Sometimes, because of the extensive use of the same crop, the soil loses its fertility power, resulting in crop failures [5].

We are trying to address these issues using the decision tree (ID3), K-means data mining algorithms, and many other data mining algorithms. Our e-portal will assist farmers in selecting the crop that is most suitable for their soil. We consider various parameters like soil color, soil moisture, soil pH value, season, rainfall, temperature all at once. Based on these parameters and using ID3 algorithm we will decide the best crop for the farmer [6].

Decision Tree is a well-known algorithm used in prediction. We have attempted to research the influence made by the decision tree in the productivity increase of crops. In this method, the first and foremost step is data collection and data filtering. First, the data set is collected, and then data filtering is done based on the conditions required by the crops for cultivation. As we are using ID3 algorithm next step is the calculation of entropy and information gain using the formula for the selection of root (test) node. Once this is done decision tree is formed and later error percentage is calculated. Testing the errors is also one of the most important aspects of applying the decision tree algorithm. An extensive collection of data set is taken, out of which 70% is used for training and rest for training. After applying ID3 algorithm to this dataset, we get a rule base. This rule base would be provided as output rules at the backend. The farmer has to select his region as input. This input follows the rule base and based on these two aspects we will select the crop. For fertilizer recommendation, the Random forest algorithm is used because of better efficiency and accuracy than other algorithms. For pesticide prediction, based on the extraction parameters, the algorithm predicts whether the crop is going to get any pest and disease attacks. We use the CART algorithm to predict the condition of the plant based on the data. ANN is used to detecting plant swelling (moisture content), burning sensation, disease, and pest along with soil analysis. Fuzzy C-means algorithm is used to identify the pest and disease present on the farm [7, 8].

In Fig. 2, shows the proposed solution. The main idea here is we have reduced to single Middle hop or agent, which helps the farmer to buy seeds, fertilizer, and pesticides and sell the crop at a better rate. We are presenting a conception of Portal. This portal is acting as the channel between the farmer and the marketing world. The overall idea behind the portal is that it is appointed for a specific area. Farmer contacts the portal manager depending upon his needs related to farming. The portal manager helps him regard the soil productivity or the essences to be adopted corresponding to


**Fig. 2** Proposed solution

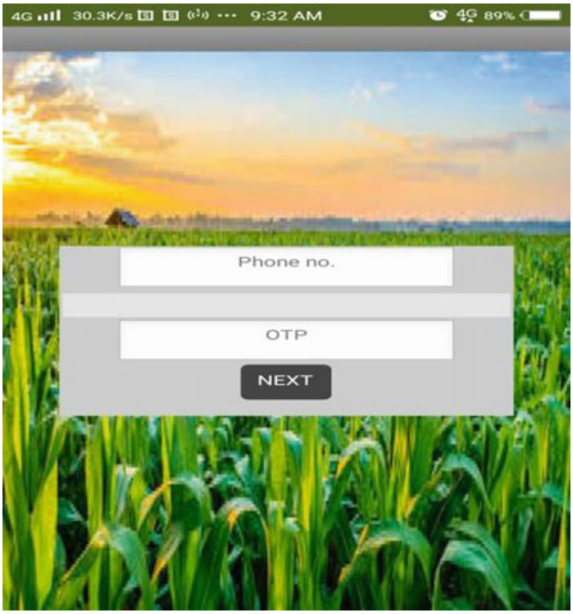



**Fig. 3** OTP generation and recognition for the given phone number

his soil types such as fertilizer, seeds, or pesticides. He provides knowledge regarding the land or crop productivity and he also helps them in buying the above-mentioned essential items with the use of our e-commerce portal.

As showed in Fig. 3, Farmers sign-up to the portal by entering phone numbers and gets verified through OTP. We are also providing the option to choose the locality and the speech input so that the user can select and operate the portal conveniently. For illustration, once the farmer logs in he have to select his locality operating the drop-down menu, he can apply speech input as well for the same purpose. After this farmer enters the main page as shown in Fig. 4, where he has to click on fertilizer. Here the farmer in result will get which crops can be grown by using decision tree and multiple linear regression ML algorithms. Later, he picks the recommended products that he prefers to purchase from the portal which is in Fig. 5. In Fig. 6, the main idea for the model is explained. Here Farmer has to enter the quantity of the product. As result, it shows the required amount to be paid and the remaining amount to be paid later. As shown in Fig. 7, the farmer should choose either a 3-months or a 6-months policy then only he will be allowed to place the order. And Farmer should complete his profile to place the order. In Fig. 8, the farmer has to provide his Aadhar card as identity proof and basic details. Now farmer will be able to place an order, once the farmer placed an order in turn it will ask for a set date and place an order as shown in Fig. 9. Once the order is established by an authenticated customer, the portal holds the authority of selling the correct and certified products to the purchaser (Fig. 10).

After the production of a crop, farmers need not go to the market for the selling of crops. They directly reach out with the same portal manager as approached before and handle over him the crop. To achieve this, the farmer has to go for crop section


**Fig. 4** Selection of the desired product

and he has to select a crop which he has to the sale as shown in Fig. 10. As shown in Fig. 11, here a farmer can send an SMS to the portal manager to notify that his crop is ready to sell. And all these transactions are maintained in DLT. In Fig. 12, we have provided a Dashboard for the farmer. The farmer can contact the portal manager through the contact counter which is in Fig. 13. The portal manager's role here is to trade the crop in the market at a reasonable cost. The non-repaid amount from the farmer is deducted from the profit earned and is handed over to the farmer. The portal manager also collects some amount of share, which is comparatively very less to the current situation. The portal manager holds the whole and sole responsibility of this e-agricultural portal. He has to be a warranted custom with a degree, competencies, prestige in that specific locality, and necessarily all the required knowledge. We also apply distributed ledger technology to keep records of the deals performed. A distributed ledger is a database of transactions that can be shared across multiple nodes in a network so that each participant has their copy, any changes done can be reflected all simultaneously. In traditional systems, ledgers are based on double-entry book-keeping wherein for every debit there must be an equal credit. When a transaction is processed, one organization records the debit, and the other records


**Fig. 5** List of fertilizers

credit. Whereas, in a centralized ledger a central authority maintains and records a single ledger to other market participants. Here, the trust is placed on the third party who maintains the centralized ledger. But, in the case of a decentralized ledger, there is no centralized data storage or authority, instead of a database that can be shared across the network irrespective of geographical or any other constraint. DLT allows multiple entities to store a copy of the ledger and compete to update the ledger independently while following the rules of the system. Participants of the network agree to specific rules for changing the ledger and the process through which all participants validate transactions or modify the ledger is called consensus algorithm. In DLT, every account has a digital public key and private key pair which is used to

Fig. 6 Details of product and price


sign and encrypt blocks, using hash functions. Here, the ownership is mentioned in a publicly shared e-ledger which is made tamper-proof through cryptography. This is the ideal case for our proposed system, even though the portal is being maintained by the portal manager, the complete control of the system is not handed over to him. All the transactions will be transparent, immutable and the whole system will be based on a consensus mechanism. The use of distributed ledger makes the portal more reliable such that no malicious activities can take place with the system.

## 4 Conclusion

In this paper, we suggested a comprehensive view of the e-agricultural system. Usually, the farmers have the mindset of planting the same crop, using more fertilizers, and following public opinions, but sometimes that may lead to loss. We can help them extensively by providing a user-friendly portal anyone can use even with little technical knowledge. As the application will be region-based, and the portal managers will also be the local people, it eliminates the issue of language. And also speech-text conversion technology is provided so that portal can be used easily. Our portal also helps in reducing the overall costs by eliminating the unnecessary middle persons which the farmers had to contact earlier to buy or sell crops. The ability

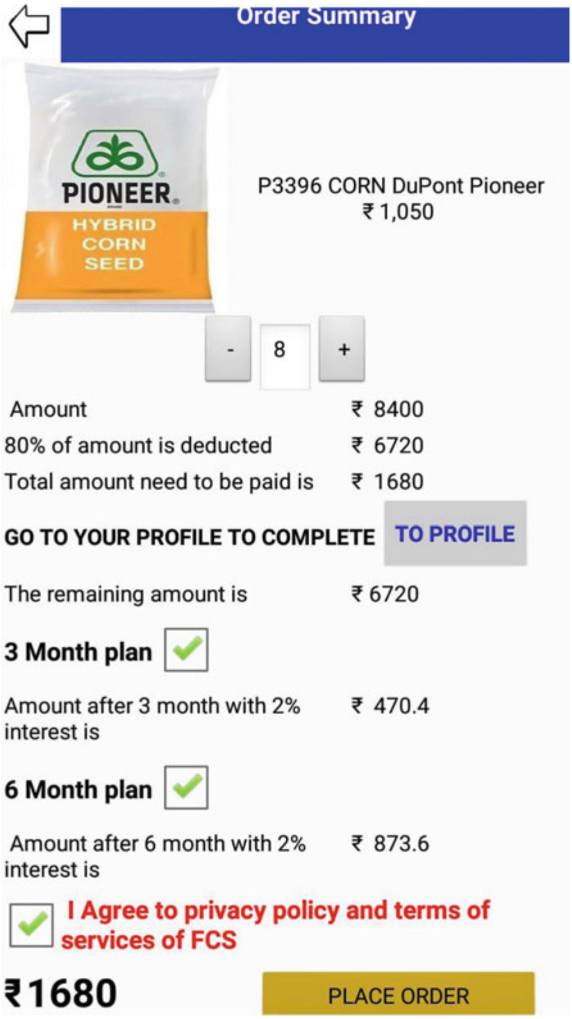



**Fig. 7** Placement of an order with payment policy description and agreement

to decide the correct products to be used in fields can help farmers in many ways, such as the overall unnecessary cost can be reduced, wastage of products can be reduced, loss done because of the usage of improper fertilizers can be eliminated, pesticides cost can be reduced. The transactions performed on the portal are more secure and reliable with the usage of Distributed Ledger Technology. DLT is a much better solution than the existing pen-paper documentation used by the farmers. The major component of our system is the prediction of soil fertility, the use of distributed ledger, providing essentials to farmers at a better rate, and helping them to earn more and more profit from their work. The proposed system takes into consideration the data related to soil, weather, and past year production and tells which are the profitable crops to be grown in his field, what fertilizers can be used, and which pesticides

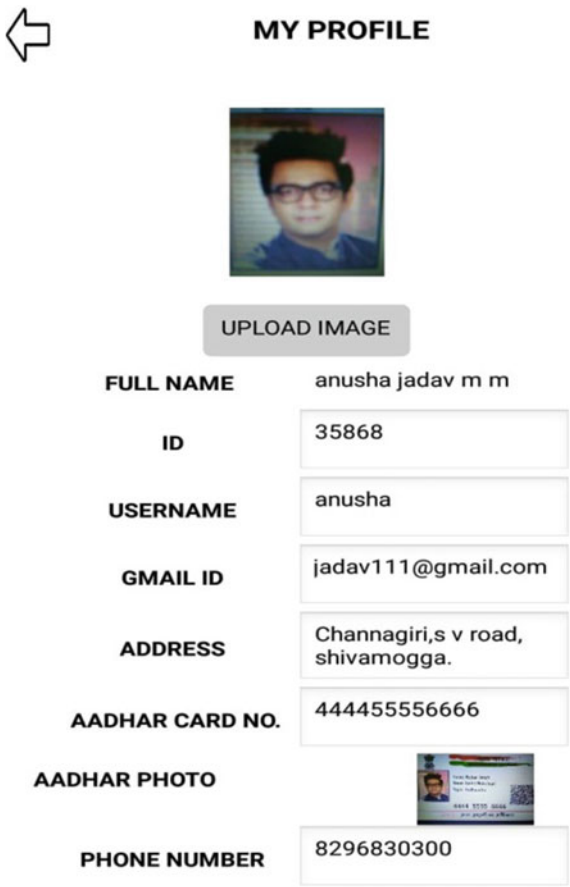



**Fig. 8** Detail and authentication of farmer

can be used. Our presented solution serves as a bridge between the farmer and the market. It eliminates all the unnecessary members, serving the farmer a better share of profit.

## 5 Future Work

As future work, we plan to make the system fully decentralized, which will intensify the Security, privacy, and reliability of the portal. All the farming devices can be connected over the internet using IoT. The sensors can be applied to the farm which collects all the real-time information about the soil and weather like the moisture, temperature, acidity, etc. We are also developing a user-friendly interface that brings a transformation in the basic techniques of crop production and realizes a qualitative leap in agricultural activity.

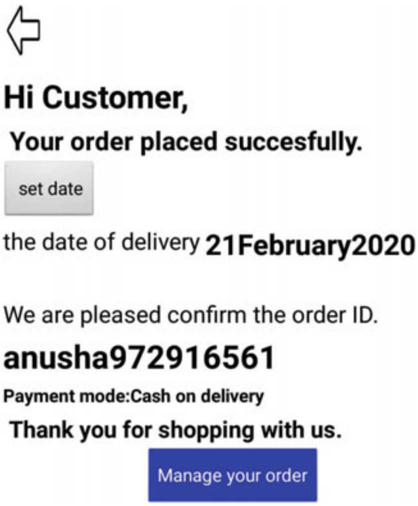



**Fig. 9** Confirmation of order placement

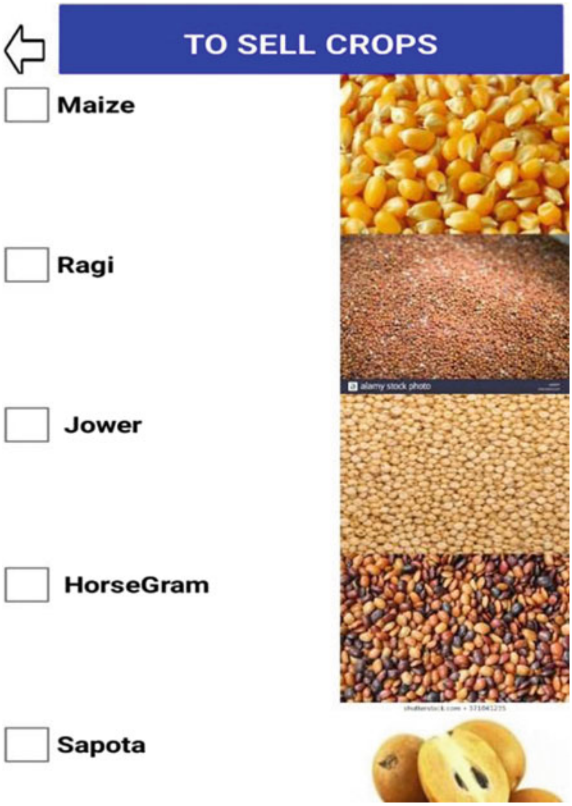



**Fig. 10** Selection of crops for sale after yield


**Fig. 11** Notify the merchant by sending an SMS

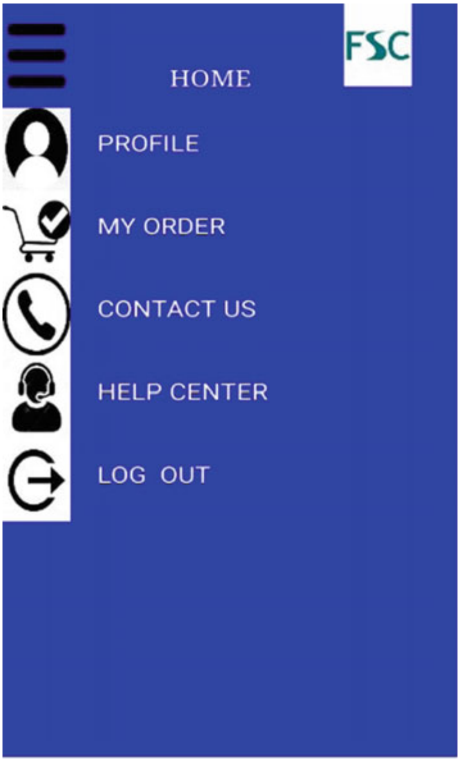



**Fig. 12** Dashboard

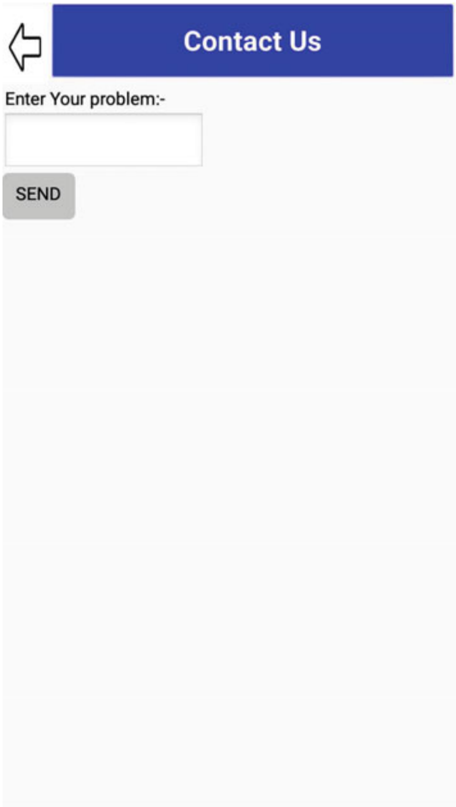



**Fig. 13** Contact counter

# References


1. Pande A, Purohit S, Jadhav S, Shah K (2019) Optimum crop prediction using data mining and machine learning techniques. Int J Res Appl Sci Eng Technol (IJRASET) 7(III). ISSN: 2321-9653
2. Bhende M, Avatade MS, Patil S, Mishra P, Prasad P, Shewalkar S (2018) Digital market: E-commerce application For farmers. In: 2018 fourth international conference on computing communication control and automation (ICCUBEA), Pune, India, 2018, pp 1–7. https://doi.org/10.1109/ICCUBEA.2018.8697615
3. Bhende M, Avatade MS, Patil S, Mishra P, Prasad P, Shewalkar S (2018) Digital market: E-commerce application for farmers. In: 2018 fourth international conference on computing communication control and automation (ICCUBEA), Pune, India, 2018, pp 1–7
4. Yi F, Rong X (2010) E-commerce applications of agricultural enterprises in guangdong province: an empirical analysis of the websites. In: 2010 2nd international conference on e-business and information system security, Wuhan (2010)
5. Zingade DS, Buchade O, Mehta N, Ghodekar S, Mehta C (2017) Crop prediction system using machine learning. Int J Adv Eng Res Dev Spec Issue on Recent Trends Data Eng 4(5)
6. Mohanraj, Naren J (2016) An architectural framework for e-agricultural system. In: 2016 3rd international conference on computing for sustainable global development (INDIACom), New Delhi, pp 1600–1604

7. Dsouza DJ, Joshi HG (2014) Development of agricultural e-commerce framework for India, a strategic approach. Int J Eng Res Appl 4(11)(Version -5):135–138. www.ijera.com ISSN: 2248-9622
8. Gupta R, Sharma PK (20185) Scope of E-commerce in agri-business in India: an overview. Int J Adv Sci Res Manage Spec (I) (2018)

# A Survey on Role of SDN in Implementing QoS in Routing in the Network

**H. Pavithra, G. N. Srinivasan, and K. S. Swarnalatha**

## 1 Introduction

QoS in Routing ensures providing preferential delivery requirements for the flow through network when it traverses through all the links. Delivery requirements can be ensuring required bandwidth, reduction of latency, reducing delay, policies, rules, packet loss reduction. Depending on the type of traffic being delivered through the network in the form of flow each application can have different QoS. In such a scenario ensuring these QoS, different for different types of applications, through traditional network requires each of the device in the path to be programmed independently. This task becomes difficult in traditional network where every device needs to be configured to achieve required QoS.

In a traditional network, every device (router/switch) needs to be programmed by the programmer who is the controller of the network and every device has to be installed with all the necessary routing information and this helps in flows traversing the network to travel between different sources and destinations. In this scenario, no single device is aware about the topology of the entire network and this also leads to another question of how to achieve QoS for a specific kind of traffic? Software Defined Networks (SDN) gives a environment in which programming the devices in the network is done from a single point of control and the hardware and the software is separated from each other and are termed as Data Plane and Control Plane where Data plane deals with only the function of forwarding the packets from one node to

H. Pavithra
Department of CSE, RVCE, Bangalore, India

G. N. Srinivasan
Department of ISE, RVCE, Bangalore, India

K. S. Swarnalatha (✉)
Department of ISE, NMIT, Bangalore, India
e-mail: swarnalatha.ks@nmit.ac.in

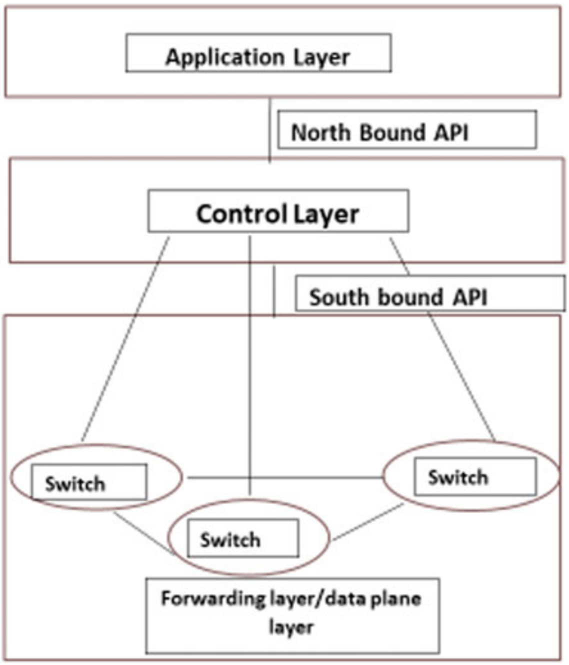
 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_31

the other by referring to the routing table which is programmed from the controller. Control plane is a software that is installed in one central system or server which is having the view of the entire network, can query every device in the network and is responsible for programming the devices with required information for performing the task of forwarding the packets from one node to the other and ensure that the packet reaches the destination.

Figure 1 shows a scenario of a SDN enabled switch where the hardware is having only data plane which does the function of only forwarding the data and all the intelligence of finding the route, configuring them is moved onto a software out of the switch called as control plane and is residing in a OpenFlow controller which is like a operating system of the network. OpenFlow controller has all the applications required to program a traffic flow and these applications can be utilized and programmed in a way required for the user by installing all the routing policies and rules and routing algorithms through which the data of an application can be forwarded through the network.

OpenFlow controller mentioned in the previous paragraph is one which is software which helps in configuring the software of the network which helps in routing. OpenFlow controller is a OpenFlow protocol enabled software which has QoS application built in it which can be programmed according to the needs of the user for satisfying the demands of the application. There are other varieties of controller. Table 1 provides the information of different types of controller and the language they can be programmed in and the applications supported. Applications listed here

**Fig. 1** SDN architecture



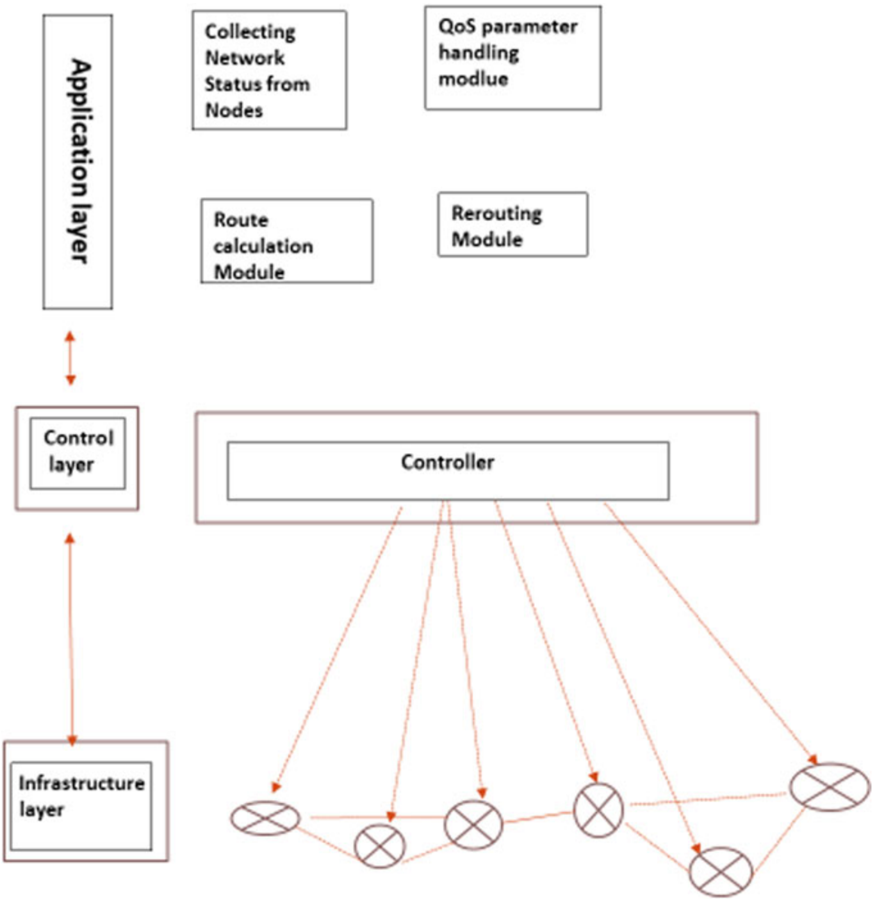



**Fig. 2** Modules of the proposed work

are in the interest of working with the proposed problem and there are few more applications supported with respect to the controllers and the sources to access these controllers are mentioned in the references.

## 2 Related Work

The concept associated with SDN (Software Defined Networks) has been existing since the time Internet has evolved, but it gained a greater significance of research interest and commercial deployment ever since 2008. SDN with OpenFlow was first introduced by McKeown et al. [1, 2]. In [3], authors have described the method of accommodating traffic engineering using SDN. Sharma et al. [4] has described the

**Table 1** Controllers-related survey information

| S. No | Controller | Language | Applications supported |
|---|---|---|---|
| 1 | Flood light controller | Java | Core applications: module manager, topology/routing manager, link discovery, packet streamer in addition to above Static flow entry pusher, circuit pusher is also provided |
| 2 | Open day light controller | Java | Host Tracker, Flow programmer, static routing, subnets, switch manager, user manager, Topology |
| 3 | POX controller | Python | Hub, switch, load balancer, firewall dump packet capture tool |
| 4 | NOX controller | C++ | Connection manager, event. dispatcher, OpenFlow manager are the core-aps along with this it also provides Net-Aps and Web-aps |

use of SDN in implementing QoS by comparison of its difficulty in implementing in traditional networks and the aspects which hinders the implementation of QoS in traditional networks. In [5] authors in their white paper have described complete routing process that takes place with SDN and have paved way for research in implementing dynamic routing within the network. This dynamic routing could also be implemented with dynamically varying QoS of the applications. In [6] authors have described the process of routing the multimedia traffic by dynamically optimizing the routes to satisfy the required QoS. Routing is normally carried out in packet switching networks where the packets are routed from one source to another destination but accommodating QoS in that routing enhances the delivery capabilities of the network and also results in end user experiencing quality especially with audio, video delivery. Delivery with QoS has been considered for quite some time in the research community nevertheless still many issues exists with it which is still ongoing as research topic. Continuing in this research area where research has been explored considering implementing QoS with routing, in [7] authors have explained the process of multipath routing in SDN, though multipath routing has been existing, deploying the multipath routing with dynamically changing application requirements is a current work taken up in this paper to be implemented.

In the next section based on the survey carried so far which is presented in the above sections a work is proposed to address the challenge of dynamically varying QoS requirements of the applications and perform routing through the network intelligently, as of now this routing method can be applied to home broadband networks and a campus network like a college and the work is being tested on simulation environment like mininet. In this scenario, few controllers were explored and tried to check which can support the coding of applications to implement the proposed work and information regarding to it is presented in Table 1.

## 3 Proposed Work

Objectives of the proposed work are as follows:

- The main objective of this research is to provide a better routing technique which can provide desired and dynamic quality of service requirements.
- To collect the network statistics efficiently
- To choose best algorithm which determines the routes
- To handle rerouting of existing flows to handle the changing QoS needs so that resources are effectively managed in the network
- To incorporate machine learning to carry out the above-mentioned objectives.

In the proposed work, the main objective of the work is to be able to collect the QoS requirements of a user dynamically. Considering a user with some data to send in a time period may have varying types of data to send and also the constraints for the same user might vary over a period of time for the same destination. considering this scenario in a campus network or for a home broadband network the objective is to be able to manage the dynamically varying QoS requirements and make a study on the frequency of the QoS parameters to be collected from flow and also handle rerouting with respect to network where a flow has been assigned a path with certain resources and that flow has less requirements of QoS provided by that path and hence reroute it in such a way that user still continues to get a seamless service and also there is an efficient management of the route within the network.

Currently, the work is being carried out in the above-proposed methodology, in the next paper detailed methodology; results will be presented as future work of this paper.

## 4 Conclusion

In the current scenario, SDN is gaining importance due to its centralized control and also the ability to provide required Quality of Service for applications separately without having to modify underlying hardware. Providing QoS to applications along with its routing is one of the important aspect as it is related to users Quality of Experience. This paper has focused on carrying a survey as what are the areas of QoS which can be addressed by SDN through its implementations. Methodology and results will be published in the next paper continuing the work done in this paper.

## References

1. Kucminski A, Al-Jawad A, Shah P, Trestian R (2017) QoS-based routing over software defined Networks. In: 2017 IEEE international symposium on broadband multimedia systems and broadcasting (BMSB), 7–9 June 2017

2. McKeown N et al (2008) OpenFlow: enabling innovation in campus networks. ACM SIGCOMM Comput Commun Rev 38(2):69
3. Agarwal, S, Kodialam M, Lakshman TV (2013) Traffic engineering in software defined networks (2013) Proceedings IEEE INFOCOM
4. Sharma S, Staessens D, Colle D, Palma D, Goncalves J, Figueiredo R, Morris D, Pickavet M, Demeester P (2014) Implementing quality of service for the software defined networking enabled future internet. In: 2014 Third European Workshop on Software-Defined Networks
5. White paper titled "demystifying routing services in software-defined networking", available in www.aricent.com/software/software-defined-networking.html
6. Egilmez HE, Tahsin Dane S, Tolga Bagci K, Murat Tekalp A (2012) KOC University, Istanbul, Turkey," OpenQoS: an OpenFlow controller design for multimedia delivery with end-to-end quality of service over software-defined networks. In: Signal & information processing association annual summit and conference (APSIPA ASC), 2012 Asia-Pacific
7. Izumi S, Edo A, Abe T, Suganuma T (2015) An adaptive multipath routing scheme based on SDN for disaster-resistant Storage Systems. In: 2015 10th international conference on broadband and wireless computing, Communication and Applications
8. www.floodlight.org
9. www.opendaylight.org

# Proficient Detection of Flash Attacks Using a Predictive Strategy

**C. U. Om Kumar and Ponsy R. K. Sathia Bhama**

## 1 Introduction

With the offshoot of Internet of Thing the number of people accessing websites has increased meteorically. This has resulted in such websites getting over loaded and causing a disruption in service. This is because the deluge of requests has caused the websites to exceed their service capabilities resulting in low performance. Such an eventuality in technical terms is known as Slashdot effect, hotspots or flash crowds [1]. This poses a serious damage to those website owners, more so if the website happens to be a commercial one, its owner will incur a severe setback in business since clients will turn to other websites expecting better performance [2]. Any website for that matter is vulnerable to such a flash crowd attack [3]. For instance, if a Government website is to release the results of a board exam or any competitive examination a number of requests would flood the website requesting for details of result of performance. Hence it is necessary for all websites to be equipped with strategies to handle a flash crowd attack [4, 5].

The availability of service through cloud computing upon the emergence of internet and wireless networking has made the security of information exposed to an even greater risk of abuse. These security threats can be categorized as malware, cracker, insider, system vulnerability, multimedia, application and service of cloud computing, zombie, social network engineering, zero-time difference attacking,

C. U. Om Kumar (✉)
Department of Computer Science & Engineering, SRM Easwari Engineering College, Chennai 600 044, India
e-mail: omkumar.cu@eec.srmrmp.edu.in

P. R. K. Sathia Bhama
Associate Professor, Department of Computer Technology, MIT Campus-Anna University, Chennai 600 044, India
e-mail: ponsy@mitindia.edu

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_32

unknown attacking and spoofing attack. Among the network attacking strategies, DoS or DDoS is disastrous as its objective is to paralyze the system.

## 2 Related Word

Shiaeles et al. [6] proposed fuzzy estimators generally work on preventing a victim from being flooded by using Inter Arrival Time between packets. This work maps the benign request distribution with the poison model. The arrival time of an attack request would not comply with that of poison distribution and hence they propose to use a fuzzy estimator to calculate mean packet arrival time for detecting a DDoS attack.

Qin et al. [7] proposed entropy based detection of DDoS traffic. They construct entropy vectors for selecting influential features in a DDoS attack traffic. They cluster all the benign traffics through the observed entropy value and identify outliers that deviate from the identified normal benign cluster.

Zekri et al. [8] proposed Machine learning based Decision tree to identify bots that generate DDoS requests. The botnet traffic consumes the available bandwidth leaving the legitimate users to wait indefinitely. The C4.5 classifiers is trained with attack signatures to effectively identify botnets raising DDoS attacks.

Nezhad et al. [9] proposed a chaos management theory to discriminate normal traffic from attack traffic. The model involves normalization followed by ARIMA model plotting. The limitation of this work is that it is confined to a single feature i.e. total packets. During flash attacks there are so many features that shows correlation among them.

Tsai et al. [10] proposed an Intrusion Prevention System (IPS) through Time Delay Neural Network (TDNN). They deploy detectors across nodes that monitor the activity and upon encountering suspicious packets the information is gathered and moved to kernal expert module which traces the source IP of the attack traffic and blacklists them.

Karimazad et al. [11] used Neural Network for detecting malicious traffic. Radial bias function captures malicious traffic through packet feature inspections. This work sniffs for TCP, ICMP and UDP packets and analyses seven feature by using Radial basis function. The RBF neural networks classify malicious traffic from normal traffic. The identified packet source IP is sent to traffic aggregation module to blacklist such similar traffic generated from distributed sources.

Deka et al. [12] proposed the study of self-similarity in attack traffic through Hurst parameter. Attack traffic exhibits identical statistical features that would repeat often as a small fractional part of a whole object. By using Hurst parameter, the normalized data can expose fractal values. These are patterns that repeat itself in a major block of a file.

Fouladi et al. [13] proposed flow based algorithms to identify malicious attacks. Enormous amount of dummy packets is despatched by the attacker to drain the victim's resources. By analysing the symmetry of the traffic distribution, kurtosis, self-similarity, skewness the model predicts features that can easily detect DDoS from benign traffic.

Robinson et al. [14] proposed ranking of supervised Machine learning algorithms through Multi criteria Decision Aid software called PROMETHEE. The objective was to reduce the Type I and Type II errors while maintaining the performance. The technique does parsing of packets to extract features, followed by data normalization. Further they proceed with classification strategies and finally rank the techniques.

Wu et al. [15] uses decision trees that classifies attack traffic through its features. The deployed classifier searches for attack pattern and if finds a hit it successfully backtracks the source IP to block the compromised device.

Jazi et al. [16] proposed to observe DoS anomalies through Cumulative Summation approach. CUSUM inspects the traffic flows to spot deviations from benign traffic. This technique is capable of identifying high/low rate Dos attacks.

Livadas et al. [17] proposed a supervised learning approach to classify IRC based botnet generated traffic. This method was tested through Naive Bayes, C4.5 and Bayesian networks by recording False Positive Rate and False Negative Rate. Though this technique detects botnet without signatures, its drawback is that it can detect only IRC botnets that follow centralized topology.

Thomas et al. [18] proposed NetBouncer. It maintains a whitelist that holds legitimate users. Packets that are forwarded to and from legitimate users/whitelist surpass the filter. When a packet arrives from a user who is blacklisted, the system sounds an alarm for intervention by the administrator. This system has a lot of practical difficulties since flash attack are raised by different geological users. Checking them and adding them to the white list degrades the performance.

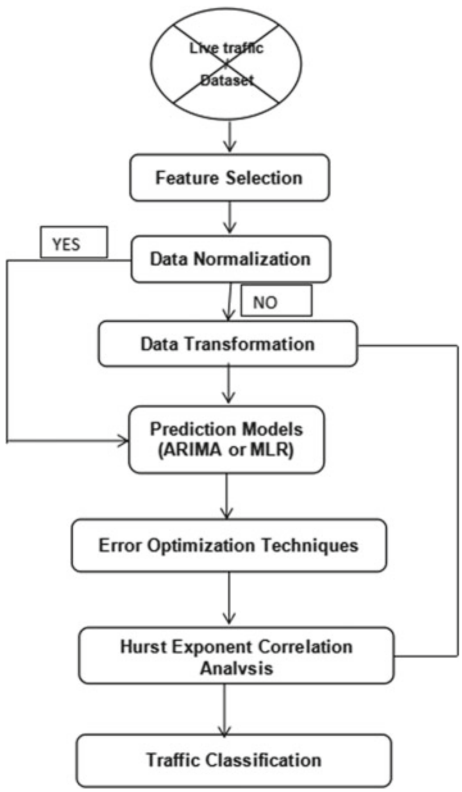
Bilge et al. [19] proposed analysing the traffic through three filters. The first filter analysed the total inbound/outbound bytes, their minimum, maximum, mean bytes over a period of time. The second filter captures the user access pattern and the final filter captures temporal behaviour to identify statistical features. They used SVM, C4.5 and Random forest techniques to detect botnets.

## 3 Methodology of the Study

The bursty nature of malicious request drains all the resources in the targeted server leaving it to become unresponsive. Identification of the influential features is necessary for the validation of the proposed prediction model. Generally, malicious data are mostly abnormal i.e. the values of the features deviate abnormally. Therefore prior to tracing the existence of a pattern, normalization of data is essential. Generally, when the relationship between any two given datasets is not linear, it is necessary to perform Transformation of data for getting a linear relationship. Performing Transformation involves Normalization and normalizing the data involves normalizing the residuals through transformation. Here Jarque Bera normalization technique is adopted. This technique's hypothesis is that if the obtained output value after normalization exceeds 0.05 it means that the data distribution is normal. On the contrary, if the obtained output value is less than 0.05, it implies that the data is abnormal. If such an inference were drawn, then the next step to be adopted is Transformation. By adopting transformation, the abnormal data gets normalized. Then the data is said to be pre-processed and is fit enough to be used by the predictive models for tapping necessary information. Some of the influential factors taken from among the standardized features for analysis are listed as Source IP, Timestamp, Flow Duration, Total Forward Packets, Total Backward Packets, Total Length of Forward Packets, Total length of backward packets, Forward Packet Length Mean, Backward Packet Length Mean, Flow Packets, Flow Inter Arrival Time Mean, Forwarded Packets/s, Packet Length Standard, Average Packet Size, Average Forward Segment Size. The procedural function of the prediction model is presented as a flowchart in Fig. 1.

This in general is the procedure for identifying the presence of a DDoS attack. The values obtained for these packets sometimes spike, at other times remain constant and sometimes give a '0' value. So the first step is to tap the influential features and extract them for identifying the presence of a trend or trace of a behavioural pattern that could be latent. Generally, malicious data are mostly abnormal. i.e. the values of the features deviate abnormally. Therefore prior to tracing the existence of a pattern, normalization of data is essential. Generally, when the relationship between

**Fig. 1** Process for detecting a DDoS attack



any two given datasets is not linear, it is necessary to perform Transformation of data for getting a linear relationship. Performing Transformation involves Normalization and normalizing the data involves normalizing the residuals through transformation.

Here Jarque Bera normalization technique is adopted. This technique's hypothesis is that if the obtained output value after normalization exceeds 0.05 it means that the data distribution is normal. On the contrary, if the obtained output value is less than 0.05, it implies that the data is abnormal. If such an inference were drawn, then the next step to be adopted is Transformation. By adopting transformation, the abnormal data gets normalized. Then the data is said to be pre-processed and is fit enough to be used by the predictive models for tapping necessary information.

Auto Regressive Integrated Moving Average (ARIMA) model by Hyndman et al. (2008) is used for analysing Time series and forecasting the traffic labels. The first step in this process is checking the stationarity of the data. Stationarity is a crucial concept in Time series analysis. It refers to the fact that the statistical properties of a process generating a Time series remains unchanged or constant over time. Next Augmented Dickey Fuller test is done for checking the stationarity of the data. A value of 0.01 is obtained which indicated that the data is not stationary. At this point ARIMA a popular automatic forecasting model is used. A univariate model is constructed by taking into consideration the ratio—Total number of packets: Each User. In addition to this, the dependency is checked among multiple independent variables by adopting the Multi Linear Regression technique. The purpose of this application is to check whether the features are influential or not. The results obtained show a certain percentage of error that needs to be optimized. So the error optimization technique was used to optimize the error. Then a Probability plotting is done to check the noise residuals and ascertain if the error follows a normal distribution pattern. Lastly, Hurst exponent is used for finding out the index of dependence and also the fractal effects.

Next Augmented Dickey Fuller test is done for checking the Stationarity of the data. A value of 0.01 is obtained which indicates that the data is not stationary. At this point ARIMA a popular automatic forecasting model is used. A univariate model is constructed by taking into consideration the ratio Total number of packets: Each User. In addition to this, the dependency is checked among multiple independent variables by adopting the Multi Linear Regression technique. The purpose of this application is to check whether the features are influential or not. The results obtained show a certain percentage of error that needs to be optimized. So the error optimization technique is used to optimize the error. Then a Probability plotting is done to check the noise residuals and ascertain if the error follows a normal distribution pattern. Lastly, Hurst exponent is used for finding out the index of dependence and also the fractal effects.

# 4 Result Analysis

The experimental analysis was performed using CAIDA dataset developed by Shiravi et al. (2012). It was a collection of a week's traffic from Monday to Friday describing the characteristics of benign traffic. Many attacks like Heartbled, web attack, infiltration, DDoS, bruteforce were generated from Tuesday to Friday. This was taken as the dataset to evaluate the performance efficiency of the proposed algorithm in this study. Out of the recorded 80 features, 15 alone were extracted for building a predictive model for DDoS detection. Figure 2 shows the distribution of pre-processed data ratio.

## *4.1 Confirmation of Normalcy of Errors*

Jarque Bera test has been used to test the normalcy of the errors using the equation

$$\mathrm{JB} = n\left[\frac{\left(\sqrt{B_1}\right)^2}{6} + \frac{(B_2 - 3)^2}{24}\right] \tag{1}$$

$B_1$ is an estimate of Skewness coefficient.
$B_2$ is an estimate of the Kurtosis measure.
$n$ is the sample size.

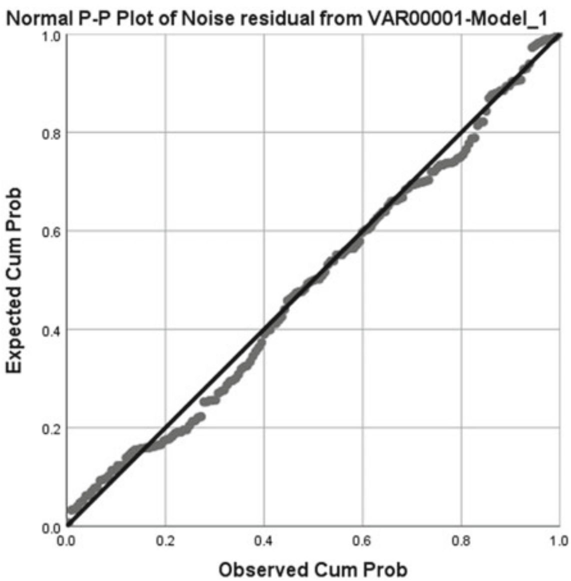



**Fig. 2** Probability plot for residuals of ARIMA (0 1 1) model

**Table 1** Auto ARIMA model function

| S no | ARIMA ($p$ $d$ $q$) model | AIC criteria |
|---|---|---|
| 1 | ARIMA(0,1,1) | 964.5897 |
| 2 | ARIMA(2,1,0) | 965.0562 |
| 3 | ARIMA(0,1,2) | 965.4285 |
| 4 | ARIMA(1,1,1) | 965.5513 |
| 5 | ARIMA(2,1,1) | 967.1439 |
| 6 | ARIMA(3,1,0) | 967.1444 |
| 7 | ARIMA(0,1,3) | 967.5199 |
| 8 | ARIMA(1,1,2) | 967.5232 |
| 9 | ARIMA(1,1,0) | 967.7219 |
| 10 | ARIMA(3,1,1) | 968.4759 |

From the result obtained ie $X$-squared = 606.66, $df$ = 2, $p$-value < 2.2$e$–16 it was noticed that the data is inconsistent and hence it needed to be transformed. So Square root transformation Sqrt $(x + 1)$ was applied.

After completing Transformation, Augmented Dickey Fuller test was carried out to check the stationarity of the data. Then it was decided to run ARIMA on the selected univariate feature. Table 1 presents the results of ARIMA models. Executing the best ARIMA model involved checking for stationarity, computing the order of differences, order of (AR) and (MA) models. They were computed by inferring through correlogram and partial correlogram. After computing $p$ $q$ $d$ coefficients for any Time Series Data the model was checked by using error optimization techniques such as $R^2$, RMSE, MSE, MAPE and AIC for selecting the best ARIMA model.

The best model among the trials was selected on the basis of criteria such as low RMSE, low MAPE, low MAE, low Normalized BIC and high $R$-Square value. Table 1 displays the top 4 models with the above mentioned criteria. Table 2 displays the top 4 models with the above mentioned criteria. Among the various ARIMA models tabled below ARIMA (011) was chosen as the most efficient working model with an AIC value of 964.5897.

**Table 2** Residual analysis of various ARIMA models

| Model | $R$-squared | RMSE | MAPE | MAE | AIC |
|---|---|---|---|---|---|
| ARIMA (1 1 0) | 0.892 | 1.975 | 67.991 | 1.461 | 967.72 |
| ARIMA (0 1 0) | 0.858 | 2.251 | 68.977 | 1.533 | 999.70 |
| ARIMA (1 1 1) | 0.908 | 1.824 | 66.523 | 1.376 | 965.55 |
| ARIMA (0 1 1) | 0.914 | 1.763 | 65.538 | 1.342 | 964.58 |
| Multiple linear regression (MLR) | 0.829 | 3.56 | 98.35 | 2.56 | 972.59 |

### 4.2 Hyndman-Khandakar Algorithm for Automatic ARIMA Modelling

This algorithm proposed by Hyndman Khandakar et al. (2008) takes into account unit root test, minimized AIC and MLE for arriving at an automatic ARIMA model. The stepwise procedure of the algorithm is given below.

Firstly, the KPSS tests are carried out repeatedly in order to find out the number of differences $0 \leq d \leq 2$.

Once the data differencing is done '$d$' times and after the minimizing of AIC and MLE, the values for $P$ and $Q$ are fixed. Instead of taking each probable combination of $P$ and $Q$ the algorithm uses the following procedure. The four following ARIMA models are fitted to start with.

ARIMA (0, $d$, 0) (0, $d$, 0),
ARIMA (2, $d$, 2) (2, $d$, 2),
ARIMA (1, $d$, 0) (1, $d$, 0),
ARIMA (0, $d$, 1) (0, $d$, 1).

Now a constant of $d = 2$ is added suppose $d \leq 1$ an additional model is fitted. The best model (i.e. the one with least AIC value) is chosen. Then the differences as observed on the current model are taken for consideration.

Now the best model, either the current model or one of those with variations becomes the new current model.

### 4.3 Akaike Information Criteria

This criterion is generally used in the selection of the best Auto regressive ARIMA model. Firstly, several probable models have to be analyzed to find out the model that could fit the data. When the obtained data is noisy generally a more complex model gives a better fit to the data (Smaller Residual Squares RSS) than a model which is not complex. Supposing RSS is to be used in the selection of the best fitting model it can be concluded that the more complex the model the better would be the fit for noisy peaks. Hence model complexity assumes importance is an important criterion in model selection. The AIC statistic is expressed using the equation

$$\text{AIC} = \ln\left[\frac{\sum_{t=1}^{n} \varepsilon_t^2}{n}\right] + \frac{2k}{n} \tag{2}$$

here '$n$' denotes the number of observations, $k$ denotes the number of estimated parameters in the model, and $\varepsilon_t^2$ represents the ordinary least square and sum of squares. The smallest AIC represents the best model among the given set of regression models.

### 4.4 Root Mean Squared Error (RMSE)

RMSE denotes a value that differentiates the predicted value from the actual value of the variables. It is considered to be standard deviation difference between the observed and the estimated value of the variable

$$\text{RMSE} = \left[\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \widehat{Y_t}\right)\right]^{1/2} \tag{3}$$

where '$Y_t$' denotes the ith observed value on dependent variable $Y$, '$\widehat{Y_t}$' is the estimated value of dependent variable and '$n$' is the total number of records.

### 4.5 Mean Absolute Percentage Error (MAPE)

The Mean Absolute Percentage error expressed in percentage is represented as follows:

$$\text{MAPE} = \sum_{t=1}^{n}\left(\frac{Y_t - \widehat{Y_t}}{Y_t}\right)\frac{100}{n} \tag{4}$$

where '$Y_t$' denotes the ith observed value on dependent variable $Y$, '$\widehat{Y_t}$' is the estimated value of dependent variable and '$n$' is the total number of records.

### 4.6 Mean Absolute Error (MAE)

The Mean Absolute Error is expressed as

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - \widehat{Y_t}\right| \tag{5}$$

where '$Y_t$' denotes the ith observed value on dependent variable $Y$, '$\widehat{Y_t}$' is the estimated value of dependent variable and '$n$' is the total number of records.

### *4.7 Coefficient of Multiple Determinations ($R^2$)*

This method checks the correctness of the statistically computed coefficient with the real data points. The $R^2$ value close to '0' indicates that the value of coefficient is far less in comparison with the real data and the value '1' indicates that the value is closely related to data points.

$$R^2 = 1 - \frac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n} \left(Y_t - \overline{Y}\right)^2} \tag{6}$$

where in '$Y_t$' denotes the tth observed value and $\overline{Y}$ Mean value of dependent variable $Y$ and $\sum_{t=1}^{n} e_t^2$ denotes the OLS residual sum of squares.

The specified variable is inspected through Probability Plot analysis. A separate line is drawn straight and the error point distribution is checked to know if the data distribution is normal. The error points lie close enough to the line which indicates that the error occurrence follows a normal distribution. Generally, some data points lie at the extreme end of the straight line, such points are called 'outliers'. These points are to be smoothened or else they would affect the statistical analysis. In some scenario, few data points will stay along the extreme ends of the straight line. These are called long tails and they exhibit patterns like more data points popping at the top of the line and continue to fall at the bottom of the line. Such long tails affect the performance of statistical procedures.

Generally, some data points lie at the extreme end of the straight line, such points are called 'outliers'. These points are to be smoothened or else they would affect the statistical analysis. In some scenario, few data points will stay along the extreme ends of the straight line. These are called long tails and they exhibit patterns like more data points popping at the top of the line and continue to fall at the bottom of the line. Such long tails affect the performance of statistical procedures. Skewed data points exhibit concave or convex curves showing data to be one sided. These errors are to be smoothened by applying equivalent power transformations. It is clear from Fig. 2 that the error points lie close enough to the line which indicates that the error occurrence follows a normal distribution.

To check the cause and effect analysis among the selected features it was decided to use the Multiple Linear Regression (MLR) strategy. Table 2 describes the coefficients of multiple variables that are considered to be dependent. In this MLR strategy the Total Forward Packets, Total Backward Packets and Backward Packet Length Mean were significant at 0.01 levels. The obtained $R$-squared value of 83% in the study shows the impact of Dependent Variables on Independent Variable. In conclusion an overview of the performance of influential features through MLR Technique is presented in the Table 3.

The graph given below shows that the predicted values and the actual values obtained overlap with each other. Figure 3 shows that the error point lies almost close with the diagonal line. Both these reaffirm the Normalcy of error distribution.

**Table 3** Residual analysis of various ARIMA models

| Parameters | Coefficient | SE of coefficient | *t*-value | *p*-value |
|---|---|---|---|---|
| Constant | 3.47 | 0.269 | 12.89 | 0.000** |
| Total fwd packets | –0.00872 | 0.00161 | –5.43 | 0.000** |
| Total bwd packets | 0.00931 | 0.00197 | 4.72 | 0.000** |
| Total length of fwd packets | 0.000007 | 2.5E–05 | 0.27 | 0.79 |
| Total length of bwd packets | 0.000001 | 1E–06 | 0.9 | 0.37 |
| Fwd packet length mean | –0.00006 | 0.00012 | –0.52 | 0.606 |
| Bwd packet length mean | 0.000004 | 1E–06 | 5.93 | 0.000** |
| *R*-Squared | 82.92% | | | |

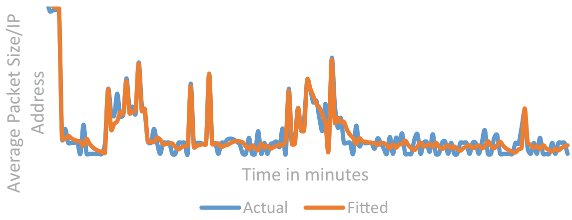



**Fig. 3** Actual versus predicted packet distribution

**Table 4** Hurst exponent estimation of univariate

| Hurst estimation to predicted Univariate feature | Value |
|---|---|
| Simple *R*/*S* Hurst estimation | 0.6678 |
| Corrected *R* over *S* Hurst exponent | 0.8543 |
| Empirical Hurst exponent | 0.6858 |
| Corrected empirical Hurst exponent | 0.697 |
| Theoretical Hurst exponent | 0.5275 |

This confirms the fact that the proposed predictive model is efficient in the detection of DDoS attack in a cloud environment.

It is a known fact that $H$ value for a DDoS traffic is always higher when compared with the $H$ value of a benign traffic. This is confirmed from the Tables 4 and 5 below

**Table 5** Hurst exponent estimation of pre-processed data

| Hurst estimation to pre-processed data feature | Value |
|---|---|
| Simple *R*/*S* Hurst estimation | 0.6585 |
| Corrected *R* over *S* Hurst exponent | 0.7935 |
| Empirical Hurst exponent | 0.6956 |
| Corrected empirical Hurst exponent | 0.707 |
| Theoretical Hurst exponent | 0.5275 |

infers the Hurst estimation of the pre-processed DDoS traffic to be 0.6585. The Hurst estimation to the univariate feature was derived as a ratio—Total Packets generated: Number of Devices the value of which seems to align with the previously calculated Hurst value reaffirming that the predictive model was highly efficient and accurate.

## 5 Conclusion

The availability of service through cloud computing has made the issue of DDoS attack more sophisticated and challenging. Effective redress to this issue is the need of the hour. The current study was undertaken with this as the objective. The empirical study was done using CAIDA 2012 dataset as sample data. The fruitful outcome obtained through statistical model is very accurate as a detection strategy which can discriminate malicious traffic amidst the benign ones and can also correctly predict the occurrence of a DDoS Flash attack. As the maxim 'Forewarned is Forearmed' goes being forearmed with the intimation of an impending attack, timely counter measures could be adopted for minimizing the disastrous consequences of a DDoS attack in a cloud environment.

## References


1. Bhandari A, Sangal AL, Kumar K (2016) Characterizing flash events and distributed denial-of-service attacks: an empirical investigation. Secur Commun Netw 9(13):2222–2239
2. Ahmad K, Abdullah E, Ramadan H, El-Hajj M, Hamieh J (2018): Performance analysis and comparison of detecting Dos attacks in Iot using machine learning, deep learning and data mining: a Survey. In: Proceedings of internet technology and secured transactions, pp 84–92
3. Prasad KM, Reddy ARM, Rao KV (2014) DoS and DDoS attacks: defense, detection and traceback mechanisms—a survey. Glob J Comput Sci Technol
4. Dhingra A, Sachdeva M (2014) Recent flash events: a Study. In: International conference on communication, computing and systems (ICCCS-2014), pp 94–99
5. Om Kumar CU, Bhama PRKS (2019) Detecting and confronting flash attacks from IoT botnets. J Supercomputing 75(12):8312–8338
6. Shiaeles SN, Katos V, Karakos AS, Papadopoulos BK (2012) Real time DDoS detection using fuzzy estimators.Comput Secur 31(6):782–790
7. Qin X, Xu T, Wang C (2015) DDoS attack detection using flow entropy and clustering technique. In: 2015 11th international conference on computational intelligence and security (CIS). IEEE, pp 412–415
8. Zekri M, El Kafhali S, Aboutabit N, Saadi Y (2017) DDoS attack detection using machine learning techniques in cloud computing environments. In: 2017 3rd international conference of cloud computing technologies and applications (CloudTech). IEEE, pp 1–7
9. Nezhad SMT, Gharavol EA (2016) A novel DoS and DDoS attacks detection algorithm using ARIMA time series model and chaotic system in computer networks. IEEE Commun Lett 20(4):700–703
10. Tsai C-L, Chang AY, Huang M-S (2010) Early warning system for DDoS attacking based on multilayer deployment of time delay neural network. In: 2010 6th international conference on intelligent information hiding and multimedia signal processing. IEEE, pp 704–707

11. Karimazad R, Faraahi A (2011) An anomaly-based method for DDoS attacks detection using RBF neural networks. Proceedings of the international conference on network and electronics engineering 11:44–48
12. Deka RK, Bhattacharyya DK (2016) Self-similarity based DDoS attack detection using Hurst parameter. Secur Commun Netw 9(17):4468–4481
13. Fouladi RF, Kayatas CE, Anarim E (2018) Statistical measures: promising features for time series based DDoS attack detection. In: Multidisciplinary digital publishing institute proceedings, vol 2, no 2, p 96
14. Robinson RRR, Thomas C (2015) Ranking of machine learning algorithms based on the performance in classifying DDoS attacks. In: 2015 IEEE recent advances in intelligent computational systems, pp 185–190
15. Wu Y-C, Tseng H-R, Yang W, Jan R-H (2011) DDoS detection and traceback with decision tree and grey relational analysis. Int J Ad Hoc Ubiquitous Comput 7(2):121–136
16. Jazi HH, Gonzalez H, Stakhanova N, Ghorbani, AA (2016) Detecting HTTP-based application layer DoS attacks on web servers in the presence of sampling. Comput Netw 121:25–36
17. Livadas C, Walsh R, Lapsley D, Strayer WT (2006) Usilng machine learning technliques to identify botnet traffic. In: 2006 Proceedings of the 31st IEEE conference on local computer networks, pp 967–974
18. Thomas R, Mark B, Johnson T, Croall J (2003) NetBouncer: client legitimacy-based high-performance DDoS filtering. In: Proceedings of the DARPA information survivability conference and exposition, vol 1. IEEE, pp 14–25
19. Bilge L, Balzarotti D, Robertson W, Kirda E, Kruegel C (2012) Disclosure: detecting botnet command and control servers through large-scale netflow analysis. In: Proceedings of the 28th annual computer security applications conference, pp 129–138

# Real-Time Image Deblurring and Super Resolution Using Convolutional Neural Networks

**Nidhi Galgali, Melita Maria Pereira, N. K. Likitha, B. R. Madhushri, E. S. Vani, and K. S. Swarnalatha**

## 1 Introduction

Digital images are electronic snapshots taken of a scene which is made up of picture elements in the grid pattern known as pixels. Each of these pixels holds important values that epitomize the intensity at that particular point. Images can act as an informative in many fields such as astronomy, medical, security systems and microscopy. Dismally, some of the images end up becoming blurry, that is as a result of the disturbance caused by external or internal factors. Many aspects can be given as a reason to obtain blurry pictures like movement of an associate object or the device, victimization long exposure times, employing a wide angle lens, etc. It is not possible to avoid the blurring of an image all the time because of which we tend to get a ruined picture. There is a way through which the initial quality of the image is often repaired. That method is Image deblurring. This aims mainly at providing sharp and useful images.

Pixels are the picture elements found in the digital image of a particular scene. Pictures are often obtained from various sources for instance, photography, astronomy, medical imaging, and remote sensing, and so on. Valuable information is found in the pictures obtained through the above-mentioned sources. While we capture an image we always wish to obtain an image which is very similar to the original scene. But mostly, the captured images would not be a replica of the scene rather they will be corrupted with blur and noise. Because of this, the main reason why the image was captured will not be fulfilled that is, the valuable information which is to be delivered will be lost. That is why image deblurring should be done to gain valuable information from the pictures. Image deblurring is a process of removing blurred

N. Galgali · M. M. Pereira · N. K. Likitha · B. R. Madhushri · E. S. Vani (✉) · K. S. Swarnalatha
Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore 560064, India
e-mail: vani.es@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_33

patches from an image. Image deblurring can be applied in everyday photography, security systems involving a camera, medical imaging and remote sensing.

Image blur has many causes such as the (a) Movement of subject: motion is one of the main causes of blur in photographs. Motion blur is caused if the subject moves while the image is being captured. Low light conditions and low shutter speed also contribute to this kind of blur. The shot appears to be blurred due to irregular movements or shutter speed. (b) Improper holding of the camera: Although this might seem insignificant but not holding the camera properly leads to foggy images. Poor handling of the image capturing device and poor technique are to be blamed for the blur caused. One needs to have a firm grip of the camera. (c) Very High ISO: When the camera is hand handled using cameras with high ISO settings avoids the blur that could be caused by shaking of hand or poor grip. But with high ISO settings noise reduction is applied to the captured image which erodes fine details. (d) Dirty lens: Despite using proper technique greasy smears, fingerprints or moisture on the camera lens may degrade the image quality to a great extent. Use of a good quality, clean lens can help in avoiding this kind of blur. (e) Wrong focusing: It is likely that the camera will lose focus if there are many objects between the lens and the subject or if the subject is not static. Light can also be a factor that affects focus. (f) Aperture of the lens: Both large and small apertures have their cons. While small apertures are better to shoot landscapes where depth of the field is in focus in large, it causes an optical effect called diffraction that blurs out essential details. This happens if the aperture is too small. If the aperture is too large the shutter speed becomes significantly low, causing blur.

The convolutional neural network (CNN) is a model that is designed to operate on two-dimensional image information. It also works with 1-d and 3-d information. The name of this network springs from the central layer in CNN. This layer performs an operation called 'convolution' which involves the multiplication of the input and an array of weights called filter or kernel.

The filter sized input and filter are multiplied to get a scalar product by element-wise multiplication as the filter is smaller than the input. A single value is obtained by summing these product values. The use of a filter smaller than the input is intentional as it allows the input array to be multiplied by the identical filter multiple times at different points on the input. The filter is also applied to every overlapping part of the filter sized patch of the input. This is a powerful idea if the filter is intended to detect a selected sort of feature in an image. The filter should be able to extract the feature from anywhere within the image.

Only one value is given as the output when the filter is multiplied with the input array once. A two-dimensional array is obtained as a result of applying the filter to the input array multiple times. This two-dimensional array is called the 'feature map'. After this feature map is made, each value within it can be passed through nonlinearity.

Convolution neural networks have changed the face of computer vision. Filtering noise in images in order to construct a high quality image is important for further processing the image for object recognition, object detection, tracking. Recently single image super resolution and deblurring have witnessed a boost in performance

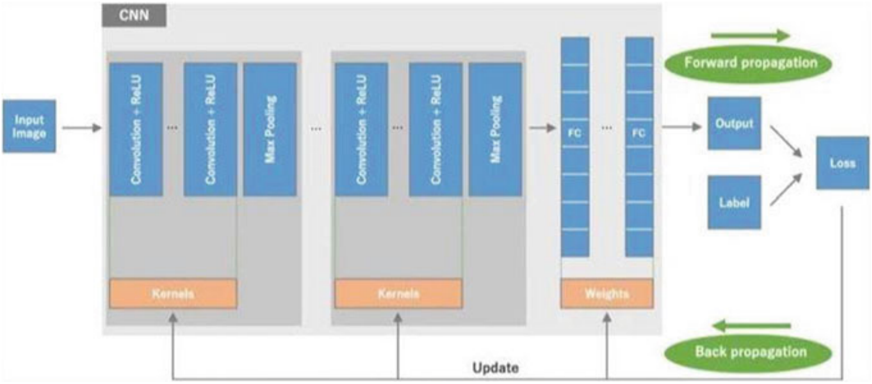



**Fig. 1** Convolutional neural network

due to CNN as it has significantly increased efficiency, accuracy, speed in parallel computation. Altering the design and training of the model also becomes easy with CNNs. The major advantage this network provides is that they link the low resolution input image to the high resolution output. This linking is made possible by learning an end-to-end mapping function between the input and output. In this paper, we use a linear CNN model that efficiently removes Gaussian noise and performs super resolution simultaneously of the image when the degradation is due to down-sampling. In this paper, we use a linear CNN model that efficiently removes Gaussian noise from the image and performs super resolution simultaneously when the degradation is due to down-sampling. Figure 1 displays the same.

## 2 Related Work

Single image super resolution aims to produce a high resolution image from single resolution one, taking approximation to a high resolution image that is unknown. Image resolution is increased to help in various fields like medical, surveillance cameras, etc. Because of the intricacy, it is difficult to solve this problem. Tsai [1] was the first one to discuss the topic of single image super resolution. To solve the problem of ambiguity in the solution the example-based approaches like the state-of-the-art methods are used. The example-based methods are categorized into two classes: internal example-based methods and external example-based methods. Most of the techniques for deblurring rely on the estimation of the blur kernel and de-convolution of degraded image to produce a sharp image. Kundur [2] gives a blind de-convolution algorithm which needs knowledge of the original image as well as the blur kernel. Kundur [3] introduces a blind de-convolution method for deblurring which requires no prior information about the kernel. Whyte [4] defines a new algorithm to find the blur kernel by using a combination of blurred images.

Yang [5] uses two images—a blurred one (B) and the other image (R) having sharper edges than B and the difference between them is used to estimate the blur kernel. The presence of significant amounts of noise affects the process of kernel estimation resulting in a distorted output image. [6] Suggests a kernel free method using a pair of blurred and noisy images and the optical flow between them to obtain the de-blurred image. But if the motion gap between the image pair is more the optical flow between them is affected resulting in a distorted image. Sun [7] makes use of CNN for non-uniform deblurring. The CNN used is meant to find the blur kernel value. [8] Introduces a blind de-convolution method by Dilip Krishnan which estimates the blur kernel by giving lowest cost to true sharp image. Levin [9] uses a MAP estimator for deblurring. From the different techniques surveyed in [10, 11] is can be concluded that it is difficult to remove average blur from a degraded image and that CNN-based methods produce a higher PSNR value than other techniques. Al-falluji [12] also presents a survey of super resolution methods and concludes that CNN-based methods are a lot more accurate.

### *2.1 Framework*

Image deblurring is the process of removing additional noise and blur from images which hinders the task of determining the original image. Images might be corrupted because of motion blur, noise and camera mis-focus. For the images that are corrupted by noise the best approach is to compensate for the degradation caused.

The workflow of the project is as shown above (Fig. 2). The Steps followed for deployment of the system are:

1. Initially, a dataset of around 90 images [5] is chosen for training and validation of the model. For the pre-processing augmentation is done to fully exploit the available dataset. Sub-images are obtained from the high resolution dataset by random cropping. So, the 90 images result in more than 21,000 training sub-images. Gaussian blur is then applied to the high resolution sub-images to produce corresponding blurred low resolution sub-images, both of which serve as the input for training and testing. Down scaling factor is used to down sample the images and bi-cubic interpolation is used for up-sampling to initial

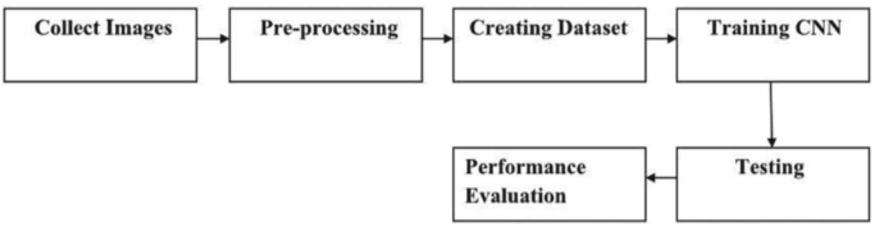



**Fig. 2** Workflow of deblurring system implementation

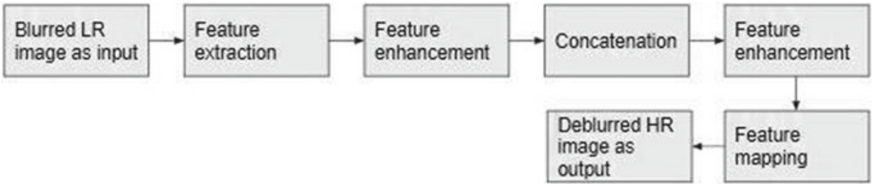



**Fig. 3** CNN architecture [13]

high resolution input size. Thus, this network produces output which is of the same size as the input.

2. For training a convolutional neural network comprising 5 layers is designed. It includes 4 convolutional layers for extracting features from the input images and enhancing them. In addition to this, a concatenation layer is used for concatenating the features extracted by convolution. The model is trained and validated on the pre-processed data to obtain the final set of updated weights.
3. The Model is Tested on Datasets 5 and 14. The Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are Calculated for Bi-Cubic as Well as the Proposed Model.

### *2.2 Methodology*

In real-time deblurring, the degraded low resolution image is given as input to a deblurring system and a sharper/clearer version of it is produced as the output. This might be applicable in the fields of astronomy, security systems and medical imaging. Figure 3 shows the detailed block diagram of the deblurring system implemented.

### *2.3 CNN Architecture*

Input Layer: this layer takes the blurred Low Resolution (LR) image '$x$' as input.

First Layer: this layer is feature extraction to compute low level features and contains 32-feature maps with filter size 9×9.

$$F_1(X) = \max(0,\ W_1 \times F_0(X) + b_1)$$

Second Layer: this is a feature enhancement layer which enhances the noisy features extracted from the previous layer and contains 32 feature maps with filters of size 5×5.

$$F_2(X) = \max(0,\ W_2 \times F_1(X) + b_2)$$

Third Layer: this layer concatenates features from the first two layers creating a merged vector of enhanced features.

$$F_{12}(X) = \text{merge}\,(F_1(X),\ F_2(X))$$

Fourth to Eighth Layer: these are the enhancement layers to further remove the noise from the output of the previous layer. As the blur level increases the enhanced features in the second layer become less prominent. Thus five layers are used since each layer has limited capability to enhance the features.

$$F_4(X) = \max\,(0,\ W_4 \times F_{12}(X) + b_4)$$

$$F_i(X) = \max(0,\ W_i \times F_{i-1}(X) + b_i)\, i \in \{5, 6, 7, 8, 9\}$$

Ninth Layer: this layer performs the final feature mapping.

Output Layer: this layer reconstructs the HR image.

$$F(X) = W_{10} \times F_9(X) + b_{10}$$

## 3 Results and Evaluation

The CNN is tested on images from set 5, set 14, Historical and Berkeley Dataset (BSDS100) that are standard datasets with different values of blur ($\sigma$). The performance evaluation is done by calculating Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) for the results of implemented system as well as the results of bicubic interpolation with respect to the ground truth.

The PSNR is the ratio of the maximum value of a signal to the value of the noise that distorts the image affecting its quality. The higher the PSNR value, the better the blurred image is reconstructed to match the original high resolution image.

The SSIM is a quality index that indicates how much a reconstructed image is similar to the original image. SSIM values lie between 0 and 1 where an SSIM value of 1 means the two images being compared are exactly same. In other words, a SSIM value closer to 1 indicates a better reconstruction of the degraded image.

$\sigma = 1$


**ORIGINAL HR**

**DEGRADED INPUT**

**BICUBIC OUTPUT (PSNR=24.24)**


**IMPLEMENTED MODEL OUTPUT**

**(PSNR=27.83)**

$\sigma = 2$

**ORIGINAL HR**

**DEGRADED INPUT**

**BICUBIC OUTPUT**

**(PSNR=20.72)**

**IMPLEMENTED MODEL OUTPUT**

**(PSNR=23.43)**

$\sigma = 3$

**ORIGINAL HR** **DEGRADED INPUT**

**BICUBIC OUTPUT** **IMPLEMENTED MODEL OUTPUT**

**(PSNR=23.37)** **(PSNR=27.24)**

$\sigma = 4$

**ORIGINAL HR** **DEGRADED INPUT**

**BICUBIC OUTPUT (PSNR=23.76)** **IMPLEMENTED MODEL OUTPUT (PSNR=25.11)**

Table 1 shows values of PSNR and SSIM for different images.

From the results, it can be deduced that the deblurring system deployed produces a higher PSNR value compared to bicubic and produces a more visually appealing output.

## 4 Conclusion and Future Scope

The paper comes up with the following results: (1) Implements an image deblurring system using deep learning techniques along with image processing techniques. (2) It is able to achieve a better Peak Signal-to-Noise Ratio in comparison with bicubic. Deep learning techniques have brought about a revolution in the field of image processing. That being said, there is a lot that is yet to be explored in the field of neural networks and deep learning. Despite the advances that have been achieved, many challenges such as very high computational cost, over-fitting/under-fitting of data, need of a large dataset, etc. exists preventing deep learning methods from completely replacing traditional image processing techniques. Extensive work and research might help in overcoming these challenges.

**Table 1** PSNR and SSIM values for images taken from different datasets

| Images | Dataset | Blur level | BICUBIC PSNR(dB) | CNN PSNR(dB) | CNN SSIM |
|---|---|---|---|---|---|
|  | SET 5 | 0.1 | 35.08 | 37.21 | 0.97 |
|  | SET 14 | 0.4 | 22.84 | 2.15 | 0.89 |
|  | SET 14 | 0.7 | 32.67 | 32.11 | 0.96 |

(continued)

**Table 1** (continued)

| Images | Dataset | Blur level | BICUBIC PSNR(dB) | CNN PSNR(dB) | CNN SSIM |
|---|---|---|---|---|---|
|  | BSDS100 | 1 | 20.94 | 22.72 | 0.80 |
|  | HISTORICAL | 1.5 | 22.23 | 26.02 | 0.84 |
|  | SET 5 | 2 | 23.94 | 29.33 | 0.91 |

(continued)

**Table 1** (continued)

| Images | Dataset | Blur level | BICUBIC PSNR(dB) | CNN PSNR(dB) | CNN SSIM |
|---|---|---|---|---|---|
|  | HISTORICAL | 2.5 | 23.48 | 26.33 | 0.79 |
|  | BSDS100 | 3 | 27.54 | 30.56 | 0.80 |
|  | BSDS100 | 3.5 | 21.69 | 25.83 | 0.72 |
|  | BSDS100 | 4 | 21.91 | 24.44 | 0.71 |

The results of the implemented model can be improved by adding enhanced layers after the concatenation layer to get better features from the merged feature maps. During training, more number of images should be taken so that the model will be able to produce optimal results for different types of images.

## References


1. Tsai R, Huang T (1984) Multiframe image restoration and registra-tion. In: Tsai RY, Huang TS (eds) Advances in computer vision and image processing, vol 1. JAI Press Inc, Stanford, pp 317–339
2. Kundur D, Hatzinakos D (1998) A novel blind deconvolution scheme for image restoration using recursive filtering. IEEE Trans Signal Process 46(2):375
3. Kundur D, Hatzinakos D (1996) Blind image deconvolution revisited. Signal Processing Magazine, IEEE 13(6):61–63
4. Whyte O, Sivic J, Zisserman A, Ponce J (2010) Non-uniform deblurring for shaken images. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), 491, 498, 13–18 June 2010
5. Yang F-W, Lin HJ, Chuang H (2018) Image deblurring , IEEE, 28 June 2018
6. Gu C, Lu X (2017) Kernel-Free image deblurring with a pair of blurred/noisy images
7. Sun J, Cao W, Xu Z, Ponce J (2015) Learning a convolutional neural network for non-uniform motion blur. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), 769–777 April 2015
8. Krishnan D, Tay T, Fergus R (2011) Blind deconvolution using a normalized sparsity measure. Computer vision and pattern recognition (CVPR). In: 2011 IEEE Conference on, Providence, RI, 2011, pp 233–240. https://doi.org/10.1109/CVPR.2011.5995521
9. Levin A, Weiss Y, Durand F, Freeman WT (2011) Understanding blind deconvolution algorithms. IEEE Trans Pattern Anal Mach Intell 33(12):2354–2367. https://doi.org/10.1109/TPAMI.2011.148
10. Singh D, Sahu MR (2013) A survey on various image deblurring techniques. Int J Adv Res Comput Commun Eng 2(12)
11. Dhole P, Chopde N (2015) A comparative approach for analysis of image restoration using image deblurring techniques. Int J Curr Eng Technol 5(2)
12. Al-falluji AAM, Youssif AAH (2017) Single image super resolution algorithm
13. Albluwi F, Krylov VA, Dahyot R (2018) Image deblurring and super-resolution using deep convolutional neural networks

# Foggy Security

**Vivek Ghosh, Bivav Raj Satyal, Vrinda G. Bhat, Nikita Srivastava, Rajesh Mudlapur, Chinmaya Nanda, M. N. Thippeswamy, and K. Venkatesh**

## 1 Introduction

The phenomenal growth in the availability of very powerful inexpensive computing engines, long range reliable low power communication technologies and the advances made in sensor technology have enabled vast sensor networks to be created and utilized for a variety of applications. This forms the backbone of technology that has created the Internet of Things (IoT). Simultaneously, maturing of the technologies associated with "Cloud Computing" has made high performance computing resources, and large size storage available fairly inexpensively. The natural evolution of these matured technologies was their integration which has resulted in IoT systems exploiting the computing power and large storage infrastructure provided by the cloud to open new vistas of applications which were unheard of earlier. The IoT intelligent sensor elements have constraints on size, weight and power (SWAP) resulting in providing only limited signal processing capability at the signal end. Consequently, a large amount of "raw" data would have to be transferred on the communication

V. Ghosh (✉) · B. R. Satyal · V. G. Bhat · N. Srivastava · M. N. Thippeswamy · K. Venkatesh
Department of CSE, NMIT, Yelahanka, Bangalore 560064, India

M. N. Thippeswamy
e-mail: thippeswamy.mn@nmit.ac.in

K. Venkatesh
e-mail: Krishnarao.venkatesh@nmit.ac.in

R. Mudlapur · C. Nanda
Unisys India Private Limited, Bangalore, India
e-mail: Rajesh.Mudlapur@au.unisys.com

C. Nanda
e-mail: Chinmaya.Nanda@in.unisys.com



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_34

channels to the cloud for further processing thus imposing huge bandwidth requirements on the communication backbone, and also resulting in indeterministic transport delays.

Naturally, the thought was to move more computing resources to a point in between the Cloud computing elements and the Edge devices so that the raw data from the Edge devices could be processed to a large extent before being transferred to the Cloud. Thus, dawned the technology of **Fog computing and communication architectures**. As data and information are valuable, there would be a need to ensure that the data/information is not distorted by malicious entities either while processing or during transmission or storage. This would mean that all intelligent entities in the chain must be security hardened.

The objective of this research is to evolve mechanisms to security harden elements in a typical Fog architecture with focus on Fog and Edge computing elements, and associated communication links: Edge to the Fog, and Fog to the Cloud. We propose to achieve this by a combination of Attribute-Based-Encryption and the Tangle mechanism. This paper specifically addresses the issue of identification of ***rogue nodes*** and proposes a mechanism of exploiting the concepts of Tangle mechanism to security harden the system. The primary objective is to develop a robust authentication and verification mechanism that would ensure all devices connected to the network are trustworthy using Peer verification of devices in the network and to use a mechanism to secure the data and vouch for its authenticity and fidelity, by periodically interrogating the nodes and monitoring their health in order to identify potentially rogue nodes.

Our contribution to the industry mainly includes a system that uses certain mechanisms to assess the integrity and authenticity of the devices along with the detection of malicious nodes in the network using Identification of Friend or Foe (IFF) protocols.

The major outcome of the paper is to evolve a scheme for identifying assets having insecure interfaces and then protecting them using appropriate mechanisms. The system also demonstrates the necessity of protocols aiming at enhancing the security of Cloud based computing and data handling resources to take the additional attack surfaces introduced by the *fog layer* also into cognizance.

The system mainly comprises of two modules which allows mutual verification of devices in the network and the detection of malicious nodes. The proposed scheme provides for lightweight mutual authentication among the edge devices and the fog servers. The involved entities make use of MQTT protocol to communicate with each other.

The central idea is to detect potentially malicious nodes by periodically monitoring various metrics of devices for instance, the memory utilization, power consumption, communication activity and the network usage information which helps in determining its authenticity. Additionally, ***attribute-based encryption (ABE)*** is employed, which allows policy-based access controls that are cryptographically enforced.

The framework also contains a resource estimation (through metric evaluation) and detection model for new IoT devices entering the network which enables dynamic real time analysis and integrated security.

## 2 Related Work

In fog computing the assurance of the trustworthiness of the nodes is a primary concern because of its dynamic nature. The security of data as well, is of concern since the paradigm of Fog Computing revolves around working with data. Thus, verifying and ensuring that the *three* parameters of security (Confidentiality, Integrity and Availability) are met is indispensable. Fog computing provides users with services such as data processing and storage facilities.

Existent literature was surveyed in order to identify the major issues that need addressing, the studied literature identified different security challenges such as authentication, privacy, data security and various types of malicious attacks were the key concerns.

Relevant research findings have been briefly discussed below.

### *2.1 Mutual Authentication*

Authentication protocols have been utilized for establishing a secure network and are considered as an entry point to any security system. Various authentication techniques have been studied and proposed in the past. Most of the traditional authentication techniques are unable to tackle the latency issues and few of them do not even meet the expectations of a basic fog computing requirement that is mutual authentication between users and fog servers [1].

Ibrahim [2] proposed an authentication protocol for fog computing where it allows any fog user or node to mutually authenticate each other without any third party being involved in the process. The constraint of this protocol is that it forces the fog nodes to store the information related to the credentials of all fog users in the same trust domain. Such solutions cannot guarantee mutual authentication due to the unreliable nature of the fog architecture.

Authors in [3] propose an authentication technique comprising three parts namely identity authentication, data integrity and data encryption to protect confidentiality, integrity and availability in a fog environment. This technique utilizes several encryption algorithms like AES and a secure hash function SHA-1 which results as a major drawback of the scheme due to the restrictions of resource constrained devices in fog computing. Thus, modified authentication schemes need to be implemented to suit the necessities of any fog environment.

Author in [4] proposes an authentication technique which addresses the uncertainty in each authentication request and tries to come up with a data driven model to handle these uncertainties during authentication. Another mutual authentication scheme proposed in [5] authenticates fog users and fog servers anonymously. This protocol uses Pseudonym Based Cryptography (PBC). The protocols in [4, 5] require a lot of computing resources which is inappropriate for the common Edge devices which have fixed capacity and low computational power. Our solution aims to resolve these issues and enhance user privacy in a fog environment.

### 2.2 Rogue Node Detection

Rogue nodes can have several definitions. In a fog environment, a rogue node is a device which tries to compromise other legitimate nodes [6]. A device which provides unusual data is considered to be a rogue node. The field of rogue fog node detection in fog computing still lacks solid research work. Most of the available proposals target the traditional computing or cloud computing environments. However, a Hidden Markov Model [HMM} based approach has been proposed by Patwary et al. [7].

They use HMM for predicting fog node behavior based on probabilities they attach to specific attack vectors. The main drawback of this scheme is its inefficiency in a dynamic fog environment where the fog nodes in the fog layer get created and deleted dynamically. They also join and leave the fog layer in a dynamic manner.

Han et al. in [8, 9], proposed a method which protects clients from getting connected to rogue access points (AP). This scheme is a measurement-based protocol. Their proposal uses the round-trip time between the Domain Name System (DNS) server and the end users to detect rogue AP present at the client side.

Al-Otaibi in [10] presented a privacy preserving, rogue vehicle node detection scheme in VANETs. The proposed scheme employs a model that uses fog nodes to perform computation needed by nearby vehicles and collaborate with other fog nodes. It also allows the detection of rogue vehicles that provide false traffic data and eliminates the incorrect data from the estimation of the traffic situation.

Based on the quantum of work done in this area it is clear that presence of rouge nodes with the sole purpose of harming the network exist and this could lead to disastrous consequences for the organization. Consequently, in order to protect the confidentiality, integrity and enhance the availability of these information gathering and processing entities it is important to improve the overall system security and integrity.

## 3 Architecture

### 3.1 Representative Architecture

The representative architecture diagram Fig. 1 depicts the various interactions between the Cloud-Fog-Edge tiers. We propose mechanisms to security harden elements in a typical Fog architecture with focus on Fog and Edge computing elements and the associated communication links: Edge to the Fog, and Fog to the Cloud and integrating these mechanisms with the Tangle scheme.

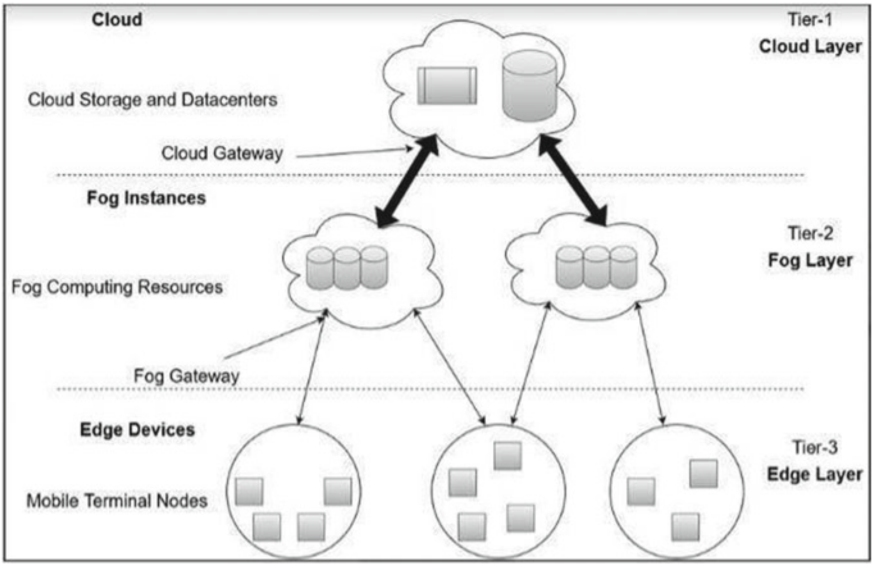



**Fig 1** Architecture of Cloud-Fog computing

The fog computing platform has a broad range of applications. The Fog processing layer several enhanced features such as rapid analysis, interoperability among devices, centralized or Machine-To-Machine management, low bandwidth consumption, low power consumption, device abstraction and many others.

### *3.2 Cloud-Fog Based Architecture*

The architecture of fog shown in the Fig. 2, represents a layered approach which includes different layers such as preprocessing and monitoring with storage and security between the transport layer and the physical layer [11].

### *3.3 Applications*

Some of the areas where fog computing proves to be of great use are:

i. Smart Home: Smart Home Applications are empowered with resources in such a way that they are integrated into a single platform using Fog Computing.
ii. Smart Grid: Micro Grid Implementation in decentralized models are deployed in the form of smart meters in electricity distribution networks which improves scalability, security, cost efficiency and rapid response.

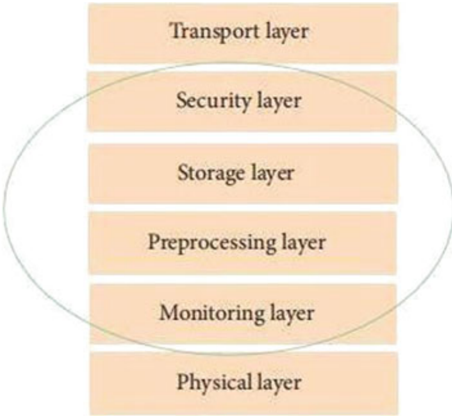



**Fig. 2** Layered architecture of fog computing

iii. Health Data Management: IoT based fog enabled cloud computing models are used in healthcare which effectively manage data of patients and diagnose the health status to identify diseases using body area sensor networks, Fog servers and Cloud Data centers.
iv. Smart Vehicles: Integration of fog computing with vehicular networks can be done with applications such as control of traffic lights, congestion mitigation, parking facility management, traffic information sharing.

As the Fog Computing platform was introduced between the end users and the Cloud, it created new areas of vulnerabilities which could be exploited for various malicious activities. For fog computing there is no specific standard of security certifications or preferred measures have been defined, unlike those available for Cloud platform [12].

# 4 Implementation

## *4.1 Methodology*

The proposed authentication scheme addresses the Cloud, the Fog server and Edge devices. The implementation is carried out using the Message Queuing Telemetry Transport (MQTT) protocol for communication, Secure Socket Layer (SSL) certificates and JSON Web Tokens (JWTs) using Raspberry Pi and the Linux systems. A rogue node detection scheme is introduced which makes use of Attribute Based Encryption (ABE) to ensure that only the devices with pre-accredited parameter values can decrypt the messages.

The building of our protocols needs these three entities:

i. Edge Devices (*E*): these are the entities (e.g. IoT devices, mobile phones) present in the edge layer, which request services from the fog layer. Each device has unique identity issued by a trusted authority.
ii. Fog Servers (*F*): these are a group of fog nodes belonging to a specific Fog zone of multiple servers under the authority of the cloud. Each F can execute the protocol to establish secure communications with its edge devices.
iii. Cloud (*C*): it is a trusted authority which is in charge of the registration and verification of the entities involved in the communication.

## *4.2 Design Considerations Mutual Authentication*

Figure 3 depicts the architectural diagram for the mutual authentication phase.

Initially, the verification is mutually completed by the fog and edge nodes using an authentication mechanism which exchanges a pair of Random numbers. Subsequent, to authentication of nodes the exchange of keys, creation of tokens and decryption of payloads is performed. Fog server decrypts contents from both the Cloud as well as the edge device and if they match, then the connection is granted for that session.

The various stages of exchange of random numbers RN between fog server *F* and the edge device is depicted. On successful verification, the Cloud C creates a token with an encrypted payload using public key of the Edge device *E*, signed using its private key. The connection is granted if the contents obtained by *F* from *C* and *E* matches.

**Rogue Node Detection**

Figure 4 depicts the architectural diagram identifying the entities involved in the detection of potential rogue nodes.

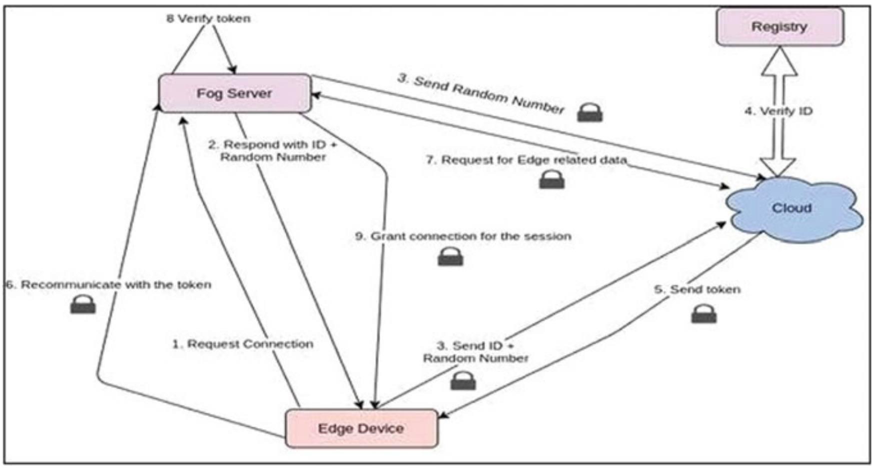



**Fig. 3** Architecture design for mutual verification phase

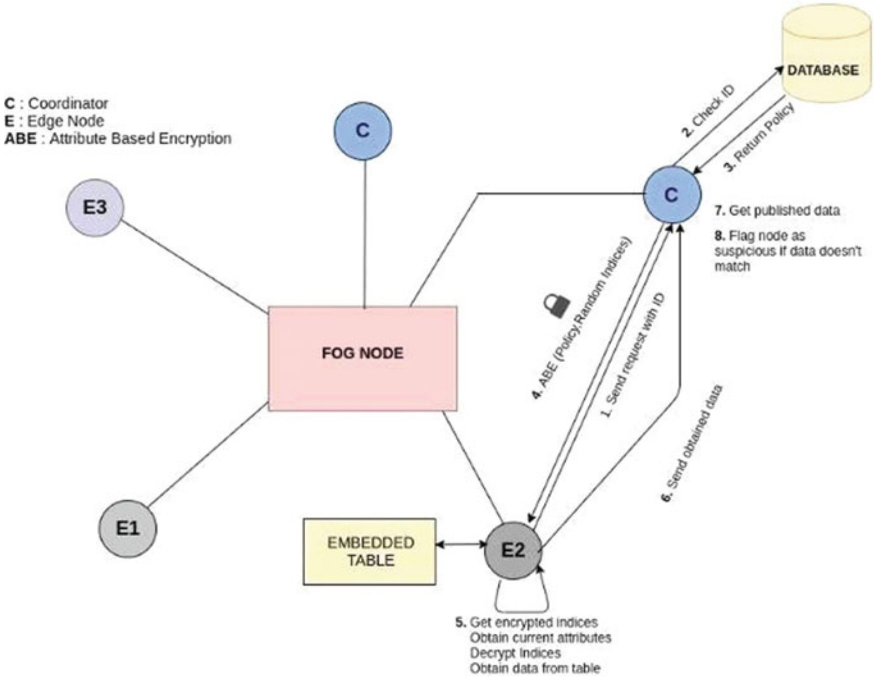



**Fig. 4** Architecture diagram for identification of rogue node phase

The Coordinator nodes are devices which have crucial information regarding the network devices and thus acts as governing agents. Embedded tables are created to capture the mapping of a set of indices to strings. They serve as a source of information the Edge devices would need to produce to validate themselves.

Initially, the Coordinator determines a certain set of indices and encrypts them such that only an entity which satisfies a certain set of parameters can decrypt that information and sends them to the edge device. If the current set of attributes in the edge device satisfy the requirements of ABE, the indices get decrypted and the corresponding data is pulled from the embedded table. If the data obtained matches the expected data, the node can be assumed to be safe, else it can be flagged as a suspicious node.

## *4.3 Description of the Process Mutual Authentication*

The proposed scheme makes use of the Message Queuing Telemetry Transport (MQTT) protocol to communicate with each other. MQTT is one of most commonly used protocols in IoT projects and its small size, low power usage, short data packets and ease of implementation made it ideal for our research. Secure Socket Layer

**Table 1** Notations used in the process of mutual authentication

| Sl No | Symbols | Description |
|---|---|---|
| 1 | REG | Registry |
| 2 | RN | Random Number |
| 3 | TOK | Token |
| 4 | KC | Public Key of *C* |
| 5 | KF | Public Key of *F* |
| 6 | KE | Public Key of *E* |
| 7 | *P* | Payload |
| 8 | IDE | ID of *E* |
| 9 | IDF | ID of *F* |

(SSL) certificates are providing us the mechanism to validate the identity of users. JSON Web Tokens (JWT) are used for secure information exchange. Table 1 refers to the notations used in the proposed scheme.

Our proposed scheme can be divided into four phases namely (Fig. 5):

i. Identity exchange phase
ii. Verification phase
iii. Mutual authentication phase
iv. Termination phase.

a. Identity exchange phase:

Before fog user registration, the exchange of identities issued to the edge device and the fog server takes place. This phase is carried out in two steps:

Step 1: This phase kicks off the communication between *E*, which makes a request and the *F* it wants to connect to. This request contains the ID of *E* that had been issued by *C*.

$$E \rightarrow F : \text{IDE} \tag{1}$$

Step 2: *F* responds to *E*, with its ID and a Random Number, RN generated for that particular session.

$$F \rightarrow \text{IDF}||\text{RN} \tag{2}$$

b. Verification phase:

Verification phase is *E* and *F* verifying themselves with *C*. The verification process is carried out between *E* and *F* with the central authority *C* as follows:

Step 1: *E* sends the IDF + RN to *C* and similarly, *F* sends IDE + RN to *C*. The parties transmit their SSL client certificate for verification by *C*.

$$E \rightarrow C : E(\text{KC}, \text{IDF}||\text{RN}) \tag{3}$$

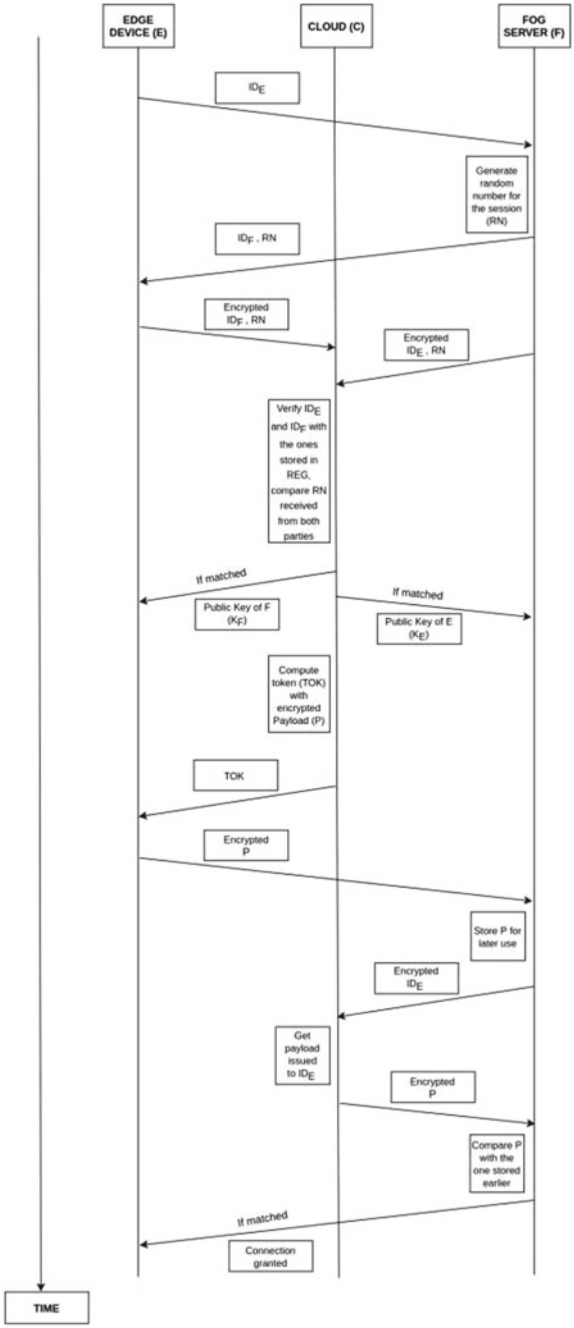



**Fig. 5** Sequence diagram for mutual authentication phase

$$F \rightarrow C : E(\mathrm{KC}\ ,\ \mathrm{IDE}||\mathrm{RN}) \tag{4}$$

Step 2: *C* verifies the IDE and IDF in its database and compares the two RNs. On successful verification, *C* sends the Public Key of *E* to F i.e. KE and of *F* to *E* i.e., KF.

$$C \rightarrow \mathrm{REG} : \mathrm{IDE}||\mathrm{IDF} \tag{5}$$

c. Mutual authentication phase:

When a verified Edge device *E* wants to access the service of a fog server *F*, *E* needs to exchange a mutual authentication request message to the *F*. Steps below shows the mutual authentication process in brief:

Step 1: *C* now creates a token whose payload is encrypted using *E*'s Public Key and *C* signs the token using its Private Key.

$$\mathrm{TOK} = E(\mathrm{KE},\ P) \tag{6}$$

$$C \rightarrow E : \mathrm{TOK} \tag{7}$$

Step 2: *E* validates the token using *C*'s Public Key and decrypts the payload using its own Private Key. Step 3: *E* now encrypts the payload contents using *F*'s Public Key and sends it to *F*.

$$E \rightarrow F : E(\mathrm{KF}, P) \tag{8}$$

d. Termination phase:

This is the concluding phase of the proposed protocol where *F* verifies the content received from *E*, with the content issued by *C* to the device with IDE. The connection is permitted if the content matches otherwise it is declined as it involves the possibility of an attack. This phase involves three steps:

Step 1: *F*, on obtaining this information from *E*, decrypts it and keeps it with itself. *F* then queries *C* about the contents issued by it to the particular IDE. *C* sends the required contents by encrypting it with *F*'s Public Key.

$$F \rightarrow C : E(\mathrm{KC}, \mathrm{IDE}) \tag{9}$$

$$C \rightarrow F : E(\mathrm{KF}, P) \tag{10}$$

Step 2: *F* decrypts the contents obtained from *C* and the previous contents from *E* obtained earlier.

Step 3: If both contents match, connection is granted for that session.

**Rogue Node Detection**

**Table 2** Notations used in the process of rogue node detection

| Sl no | Symbols | Description |
|---|---|---|
| 1 | ET | Embedded Table |
| 2 | DB | Database |
| 3 | KE | Public Key of E |
| 4 | KC | Public Key of C |
| 5 | ABE | Attribute Based Encryption |
| 6 | RI | Random Indices |
| 7 | P | Policies |
| 8 | D | Response Data |

In this section, we introduce a technique in order to provide a consistent method to identify potential rogue nodes in fog computing, which could expose the fog user devices and sensitive data of users. Table 2 lists the notations used in the proposed scheme.

The following key aspects which are central to the proposed scheme are introduced before, the scheme itself is described in detail.

i. The central idea is that every device will be represented by certain metrics that can be observed/measured and can help in determining if an unauthorized program is running on that particular device.
   Preliminary set of quantifiable metrics is:
   a. Memory Usage
   b. Network Information
ii. Attribute Based Encryption (ABE) is an approach wherein an encrypted message can be decrypted only by a party which has displayed certain pre-accredited behavioral pattern.
iii. Devices in the network are governed by a coordinator which has certain private information regarding the devices and a table of random numbers issued to them during setup.
iv. A distributed ledger (Tangle) is used to ensure that the information that is transferred between various participating entities is tamper proof.

Figure 6 depicts the sequence diagram for rogue node detection phase.

The detailed process of our proposed scheme is discussed below:

Step 1: The Coordinator, $C$ determines a set of indices and encrypts them such that an entity only satisfying a certain set of parameters can decrypt that information. The encrypted indices are denoted by ABE($m$) where $m$ = Random Indices.

$$E \rightarrow C : E(\mathrm{KC}, \mathrm{IDE}) \tag{11}$$

$$C \rightarrow \mathrm{DB} : \mathrm{IDE} \tag{12}$$

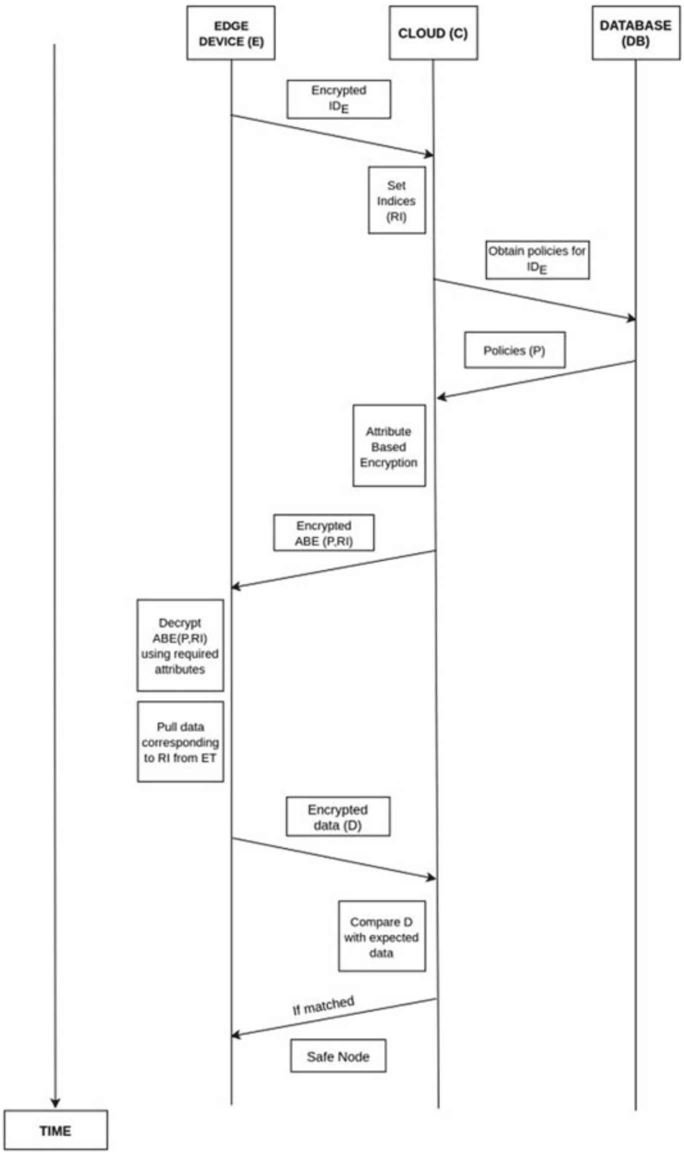



**Fig. 6** Sequence diagram for rogue node detection phase

$$\mathrm{DB} \rightarrow P \tag{13}$$

Step 2: $C$ sends ABE($m$) to the Edge Device, $E$ through a secure channel.

$$C \rightarrow E : E(\mathrm{KE}, \mathrm{ABE}(P, \mathrm{RI})) \tag{14}$$

Step 3: $E$ obtains ABE($m$).

Step 4: $E$ then obtains the set of attributes associated with it that are required to decrypt the indices.

Step 5: If the attributes satisfy the requirements of ABE($m$), the indices get decrypted.

Step 6: The data corresponding to these indices is pulled from the table.

$$E \rightarrow \mathrm{ET}: \mathrm{RI} \tag{15}$$

$$\mathrm{ET} \rightarrow D \tag{16}$$

Step 7: The data is sent back to $C$.

$$E \rightarrow C : E(\mathrm{KC}, D) \tag{17}$$

Step 8: $C$ gets the data.

Step 9: If the data obtained matches the expected data, the node can be assumed to be safe, else it can be flagged as a suspicious node.

**Assumptions** The following assumptions are made about the proposed protocol:

i. A node should be tagged as being infected only after a series of tests are performed to ensure a greater degree ofcertainty.
ii. Multiple coordinators should be used to eliminate the possibility of a single point of failure.
iii. The random indices should not be absolute but relative to a seed to reduce predictability.

## 5 Experimentation and Results

The entire implementation was carried out using Linux Systems as the Cloud Server and the Fog Server and Raspberry Pi as the Edge Device. The stated mechanisms of Mutual Authentication and Rogue Node Detection were carried out using the defined test bed. This section aims to define the outcomes of the implementation taking into account a physically simulated environment.

### *5.1 Mutual Authentication*

i. Each Fog Server and Edge Device were initially registered with the Cloud. Here the entities were given a unique identifier and added to the database. The Cloud, being the Certification Authority, was used to generate a set of SSL

Certificates for each entity and were then stored on the systems themselves. MQTT was used as the communication protocol and each of the devices had been set up to be able to carry out both publish and subscribe operations.

ii. Making use of the lightweight MQTT protocol and SSL certificates, an Edge Device and a Fog Server were successfully able to authenticate themselves with the intervention of the trusted authority, Cloud. This protocol ensures that in the fast-paced paradigm of Fog Computing, a trustworthy connection can be established between an Edge Device and a Fog Server.

iii. The authentication process needs to be carried out from start to finish without communication breakages in between. This ensured that any obsolete information could not be reused, and exploitation of gaps would not occur. The process would need to be restarted in case of any failure in the communication channel during this phase.

## *5.2 Rogue Node Detection*

i. Here a trusted entity known as the Coordinator (Another Linux System) is assigned the responsibility of identifying the potential rogue devices. The mechanism requires the Coordinator to maintain the parameters to be observed in each device and the acceptable range of values of the same. Using the parameters, the encryption policy is created which, as described earlier, encrypts the random indices.

ii. The parameters considered here are the Average Memory Utilization and Whitelisted IP Addresses the device must communicate with.

iii. The Edge Device must be set up to execute the required processes in the background which generate data with respect to the determined parameters. The generated data is used as attributes which basically act as the key to decrypt the information sent by the Coordinator.

**Outcome**

We shall look into two cases and observe what the outcomes were in each case.

i. Device functions in a non-compromised state:

A device which had not been compromised was tested by the coordinator. The following table illustrates the parameters observed.

Here it can be seen that the attributes generated satisfied the policies used to encrypt the random indices and the Edge Device was able to generate the desired response as expected by the coordinator (Table 3).

ii. Device functions in a compromised state:

In this case, a sniffer attack was carried out which required a promiscuous function to be set up on one of the Edge Devices. The same device was selected, and the test was carried out. The following Table 4 illustrates the parameters observed.

**Table 3** Parameters observed during non-compromised state of the device

| | Policy creation by coordinator | Attribute generation by edge device |
|---|---|---|
| Average memory utilisation (in percentage)<br>Whitelisted IP addresses | 25–30<br>172.17.10.242, 172.17.9.68 | 27.3<br>172.17.10.242,172.17.9.68 |

**Table 4** Parameters observed during compromised state of the device

| | Policy creation by coordinator | Attribute generation by edge device |
|---|---|---|
| Average memory utilisation (in percentage)<br>Whitelisted IP addresses | 25–30<br>172.17.10.242,172.17.9.68 | 34.1<br>172.17.10.242,172.17.9.68,172.17.11.14 |

It can be clearly observed that the Average Memory Utilization obtained was much higher than the expected range. Also, one IP address registered which was not whitelisted. The device here was not able to generate the expected response and the Coordinator flagged the device as potentially rogue for this test.

## 6 Conclusion

Security and privacy issues are well studied in cloud, fog and edge computing paradigms. The attack surface in various domains have increased due to the development and the massive deployment of edge devices. The severe attacks on edge devices can cause extensive damages to human lives and financial interests depending on the application.

In the current system, we have implemented a robust mutual verification protocol of parties when a device enters a network using MQTT Protocol where the Cloud is used as the trusted authority. The potential rogue nodes are detected using the variance in the values of the specific metrics associated with the device's current mode of operation.

The current system faces some limitations in certain situations when metric fluctuation independent activities are executed in the edge devices. Future work can be done to identify more significant attributes and refine metrics to ensure a greater degree of detectability. To address the scalability issues, multiple coordinators can be deployed for balancing the workload.

## 7 Future Work

In this work, we addressed the problem of mutual authentication of devices entering a network and the identification of rogue nodes. Our main contribution is to express this task as a optimization problem with assumptions and constraints and propose suitable methodologies.

Different kinds of adaptations, tests and experiments have been left for the future. Future work is concerned with a deeper analysis of certain techniques along with new proposals to execute different methods to introduce newer and effective technologies in securing the Cloud.

It could be interesting to consider Tangle to ensure that the information that is transferred between various participating entities is tamper proof. The Tangle protocol proposed by IOTA is currently not yet widely used and is considered to be in the transition period or the beta phase prior to eventual large-scale deployment and standardization. The Coordinator node is essentially the training wheels for the network until the Tangle can be left to evolve unassisted.

The use of Distributed Ledger Technologies like Tangle can be introduced to leverage the benefits like immutability. The nodes which have been identified as trustworthy can have their data transactions recorded into the immutable ledger periodically. These enhancements to the current scheme are proposed to be implemented in the second phase of the research activity.

## References

1. Sethi P, Sarangi SR (2017) A review article on Internet of Things: architectures, protocols, and applications. J Electr Comput Eng
2. Al-Doghman F, Zenon C, Ajayan AR, Klempous R (2016) A review on fog computing technology. IEEE International Conference
3. Amor AB, Abid M, Meddeb A (2019) SAMAFog: service-aware mutual authentication fog-based protocol. In: IEEE, 2019
4. Ibrahim MH (2016) Octopus: an edge-fog mutual authentication scheme. Int J Netw Secur 18
5. Hu P, Ning H, Qiu T, Song H, Wang Y, Yao X (2018) IEEEE Internet of Things
6. Heydari M, Mylonas A, Katos V (2019) Uncertainty-aware authentication model for fog computing in IoT. In: 2019 fourth international conference on fog and mobile edge computing (FMEC)
7. Amor AB, Abid M, Meddeb A (2017) A privacy preserving authentication scheme in an edge-fog environment. In: 2017 IEEE/ACS 14th international conference on computer systems and applications (AICCSA), IEEE, 1225–1231
8. Yi S, Qin Z, Li Q (2016) Security and privacy issues of fog computing: a survey. In: International conference on wireless algorithms and systems.
9. Patwary AA-N, Hossain N (2019) A detection approach for finding rogue fog node. In: Am J Eng Res (AJER)
10. Han H et al (2009) A measurement based rogue AP detection scheme. IEEE INFOCOM 2009. IEEE

11. Han H et al (2019) A timing-based scheme for rogue AP detection. IEEE Trans Parallel distrib Syst
12. Al-Otaibi B, Al-Nabhan N Privacy-Preserving vehicular rogue node detection scheme for fog computing. Sensors Article

# Predicting the Rate of Transmission of Viral Diseases Using GARCH

**Varun Totakura, S. G. K. Abhishek, Sangeeta Adike, Madhu Sake, and E. Madhusudhana Reddy**

## 1 Introduction

The pandemic disease COVID-19 is out spreading rapidly throughout the world. As per the statistics in April 2020, the number of cases reached around 3.27 million and the number of deaths has reached around 2.3 lakhs. In India, the number of cases is above 2 lakhs and the number of deceased people is above 10,000 and it is still growing. This disease has started in 2019 at Wuhan, China, and had spread in almost all countries on the globe. At places where people are strictly following the rules and preventive measures imposed by the respective state governments, the rate of spread of COVID-19 is falling down or very less, but at places where the rules are broken, the rate is very high. But there is a chance of having an economic crisis, and poor people will not be able to buy their essential goods. In keeping all these conditions as a priority, there is a need for an analysis to predict the impact of the virus on the people which helps to spread the awareness among people in many ways so that the people start judging the situation in a right way and the government can implement or suspend the rules or preventive measures that are taken against the spread.

The proposed statistical methodology will take the time series data [17] which contains statistical information of the number of cases, recovered, deaths in each state in India as per the records on June 15, 2020, as shown in Fig. 1. The data has been collected after taking the precautionary measures and citation. The data which we took was daily data from March 15, 2020, to June 15, 2020, which is sufficient for the prediction and can be added accuracy. The time series data will provide an opportunity to forecast future values. Based on the previous time series data and by using the forecasting statistical algorithms, the values of the same parameters can be predicted for the future. But the model which was used should be accurate so

V. Totakura (✉) · S. G. K. Abhishek · S. Adike · M. Sake · E. M. Reddy
Department of Computer Science and Engineering, Guru Nanak Institutions Technical Campus, Hyderabad, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_35

| | Date | Status | TT | AN | AP | AR | AS | BR | CH | CT | ... | PB | RJ | SK | TN | TG | TR | UP | UT | WB | Unnamed: 40 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 2020-03-14 | Confirmed | 81 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ... | 1 | 3 | 0 | 1 | 1 | 0 | 12 | 0 | 0 | 0 |
| 1 | 2020-03-15 | Confirmed | 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 1 | 0 | 0 | 2 | 0 | 1 | 0 | 0 | 0 |
| 2 | 2020-03-16 | Confirmed | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 3 | 2020-03-17 | Confirmed | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 1 | 0 |
| 4 | 2020-03-18 | Confirmed | 37 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 1 | 3 | 0 | 1 | 8 | 0 | 2 | 1 | 0 | 0 |

**Fig. 1** Data used for the proposed model

that there are no wrong decisions made. Forecasting is a tremendous task which is generally performed by higher officials of an organization for the growth of their business and even by the stock market investors to get good amount of profits. The future trend of stocks or the sales will help investors or owners to make the right decisions. The same can be used for the prediction of the impact of COVID-19 in India. We have used GARCH model in our paper for the forecasting of impact of COVID-19 as it is one of the most used method specifically for time series data. Forecasting a time series can be broadly divided into two types. If the previous values of the time series data are used to predict its future values, then it is called univariate time series forecasting. And if we consider predictors other than the series to forecast, it is called multi-variate time series forecasting. GARCH is a forecasting algorithm based on the idea that the information from past values of the time series can alone be used to predict the future values. And to confirm that the predicted data is accurate, various error calculating methods are used.

## 2 Related Work

The ARIMA model was used in the prediction of COVID-19 cases [1] in which it was mentioned that the ARIMA models can be used as an immediate tool for the prediction of time series data for health monitoring systems. It is a good model for short-term forecasting, but there should be a good procedure in the process of interpretation. The ARIMA interpreted that there will be around 200,000 cases in Italy. The usage of the ARIMA model is wide ranged. It can be used on the time series data for the prediction of daily or monthly or even yearly average. A similar kind of work was performed by [4] on the prediction of the banking stock market data. They have mentioned that the ARIMA model was very useful for the short-term analysis with few time series of data. They have used MSE method in calculation of the error value to calculate the accuracy of the model. A graph was plotted for the dataset which tells about the comparison of the banking and index. The same kind of model was developed for the Nigerian Stock Exchange by [6, 12]. The prediction of global solar radiation, by [5] in the prediction of the stock market data using EMD-HW bagging and [13] have developed a LSTM neural network model for the stock market prediction.

The generalized review about the ARIMA model for the prediction or forecasting of time series data was represented in a paper published by [8]. They have proposed a novel online method using the ARIMA model. They have theoretically proved that their method has produced better results from previous fixed ARIMA methods. They have even compared their algorithm with the previous algorithms on ARMA and have mentioned that their model performed well. A review work on forecasting the electricity price was performed by [2]. The paper will give the information about the electric price forecasting study and also interprets on the directions of the price for a decade. A generalized forecasting method was given by [9]. The first system that they have made to forecast was performed on the data from Facebook. They have used a modular regression model for their first method. And secondly, they have performed a study on tracking the forecasting accuracy of the model. A similar kind of methodology was also proposed by J. Scott Armstrong, 2001 and also by [3] in prediction of sugarcane production in India [10] The prediction of the tourism was performed using the ARIMA model. The paper has interpreted that the tourism in F. Y. R. Macedonia will play a major role in contribution for the country's economy. They have used ARIMA (1, 1, 1) model for the prediction and analysis of international tourism at F. Y. R. Macedonia. The accuracy of the model seems to be good but not perfect as there should be an increase in the accuracy for accurate prediction which will help in making correct decisions for both tourism and economic growth.

## 3 Methodology

Generalized autoregressive conditional heteroskedasticity (GARCH) is a statistical model specifically used for time series analysis. It is a generalized model of autoregressive moving average model with an extra component that is seasonal variant. Basically, these two models are used for understanding and forecasting of the time series data. GARCH is also mostly applied on non-stationary data. Because, it has an integral part which will help in removing non-stationarity of the data. A stationary time series data is the data in which the vales of the data will not depend on time. The time series which exhibits trends or seasonality are considered to be non-stationary. GARCH model is the combination of both autoregressive (AR) and moving average (MA) model. The following Eqs. 1 and 2 is the formula of GARCH ($p$, $d$). Similar to ARIMA, the least AIC value for the best $p$, $q$ value is taken and the prediction can be done.

$$\varepsilon_t = \sigma_t \omega_t \tag{1}$$

$$\sigma^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{2}$$

The term "autoregressive" in GARCH gives the interpretation that it is a linear regression model that uses its own lags as predictors. Generally, the linear regression

```
ARIMA(1, 1, 1)x(0, 0, 1, 12)12 - AIC:1217.0176398667315
ARIMA(1, 1, 1)x(0, 1, 0, 12)12 - AIC:291.8271786523849
ARIMA(1, 1, 1)x(0, 1, 1, 12)12 - AIC:174.77358103778263
```

**Fig. 2** Examples of the AIC values obtained by GARCH model

**Table 1** Table showing the results of GARCH model

| | Coef | Std err | *z* | *P* > \|*z*\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| ar.L1 | 0.2576 | 1.165 | 0.221 | 0.825 | –2.025 | 2.540 |
| ma.L1 | –0.9321 | 1.169 | –0.797 | 0.425 | –3.224 | 1.359 |
| ma.S.L12 | –0.7888 | 1.660 | –0.475 | 0.635 | –4.043 | 2.466 |
| sigma2 | 22.6742 | 29.265 | 0.775 | 0.438 | –34.685 | 80.633 |

models will work best when there is no correlation with the predictors and also when they are independent to each other. To make the time series data stationary, the most common approach which is used is the differentiating method, that is, subtract the previous value from the current value. Sometimes, depending on the complexity of the time series data, more than one differencing is performed. The example of the AIC values which are obtained by the GARCH model when fitted with the COVID-19 "Andhra Pradesh" State data are as shown in Fig. 2. Among the obtained values, the least AIC value is taken and the corresponding parameters are given to the parameter list of the GARCH model.

After obtaining the AIC values, the GARCH model is fitted or trained with the time series data which is the state COVID-19 data. And the results of the model can be represented as shown in Table 1 which contains the vales of the autoregression, moving averages and the corresponding coefficient values, standard errors, and many more. These values can be represented in the graphical mode as shown in Fig. 3 which consists of a combination of four graphs which represent the values of the standard residual, histogram plus estimated density, normal Q–Q, and correlogram graphs. These two figures contain the results that were obtained by training or fitting the GARCH model with the data from the state Jharkhand, India.

The predictions of the proposed statistical model that is GARCH model for each state can be seen in the results discussion section of this paper. Along with these, details of the predicted date from which the impact of COVID-19 may decrease is also displayed using a table. And, the error calculation and accuracy of the model are also displayed in the same results section.

## 4 Comparative Study

The comparison of prediction by the proposed GARCH model can be performed on the predictions performed by the ARMA model. The ARMA prediction model was

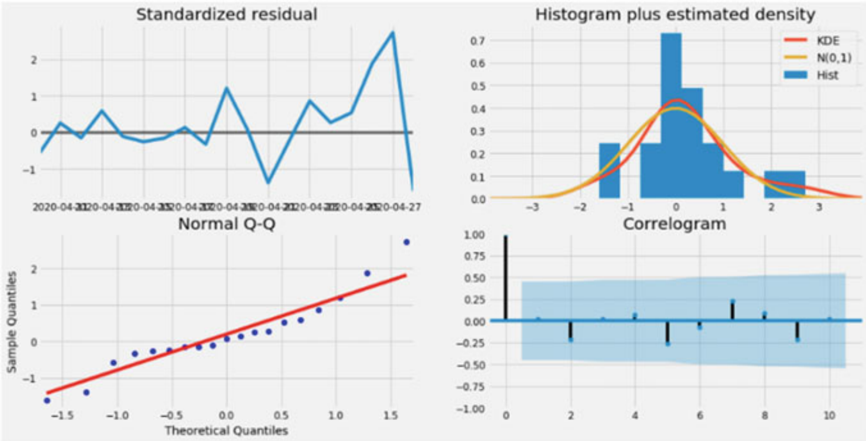



**Fig. 3** Graphs which shows the result of GARCH model

applied to predict the rainfall by [15]. They have taken rainfall data of the Republic of Indonesia from 2001 to 2013 month-wise data and fitted to a univariate time series ARMA model. In their study, their proposed ARMA model had produced results which were not synchronizing with the previous data. A study of rainfall was performed by [16] where a method for forecasting of rainfall for a year was developed using ARIMA model on the Indonesia data. The data which was used by the author for training the model ranged from 2011 to 2014. The GARCH model presented in our paper had produced better results when compared with the above-mentioned cases. The model was not over fitted or under fitted, and the usage of GARCH will reduce the time consumed to make the input data stationary as the differential part was integrated in the GARCH model, and also by using the seasonal component for the ARIMA, the results can be more accurate. Table 2 describes as easier way of comparative study for the above-mentioned methods. It gives the drawbacks of the existing models (ARMA and ARIMA) which are rectified in proposed GARCH model which is the reason behind the proposed model performing better for all types of data.

**Table 2** Comparative study with previous methods

| Paper | Model | Drawbacks |
|---|---|---|
| [15] | ARMA | Delay due to manual differentiation to make input data stationary |
| [16] | ARIMA | The model was under fitted and had produced highly normalized results |

## 5 Results and Discussion

The proposed GARCH model has produced good results with the selected *p*, *q,* and *d* values. But, as the accuracy of the model depends on the data that is given, for every state data that is given to the model, its respective *p* and *q* values are used according to the least AIC value. The graph obtained using the predicted values of the number of cases and actual values by the proposed statistical model on the time series data of Jammu & Kashmir state was displayed in Fig. 4. It proves that the model was finely fitted to that data as the predicted and the actual data seems to be equal. In the graph displayed below the blue line tells the actual or observed values, whereas the red lines give the information of the predicted information.

The prediction of number of recoveries by the proposed model in the state of Kerala using the data of the number of recoveries from that state is as shown in Fig. 5.

After the prediction of graphs, the data of the prediction produced by the GARCH model is taken and interpreted to produce the estimated date from which the number of cases will reduce gradually is displayed in the Table 3. The table consists of some states on which the model was trained by the respective state time series data along with their estimated date of recovery. The date which was shown in the model tells about the day from which COVID-19 cases will reduce gradually in that state.

In Table 3, all the other states' data predictions have given RMSE values in the range of 4–9, but Bihar, Maharashtra, and West Bengal have shown RMSE values in the range of 20–30. So, the predicted information of those states is not as accurate and can change. The accuracy of the proposed model was predicted using the produced RMSE values. For RMSE value which were less than 5, the accuracy was 95% and above, and for the values less than 10, it was 90% and above. But those states with

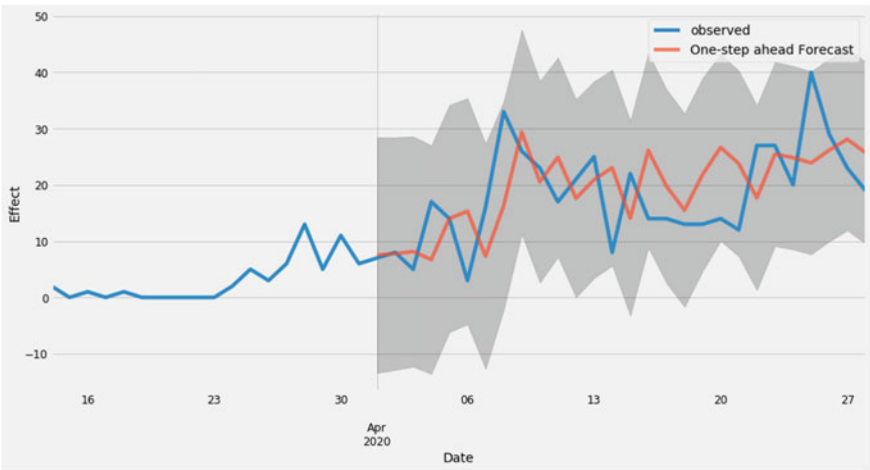



**Fig. 4** Prediction of number of cases in Jammu & Kashmir state

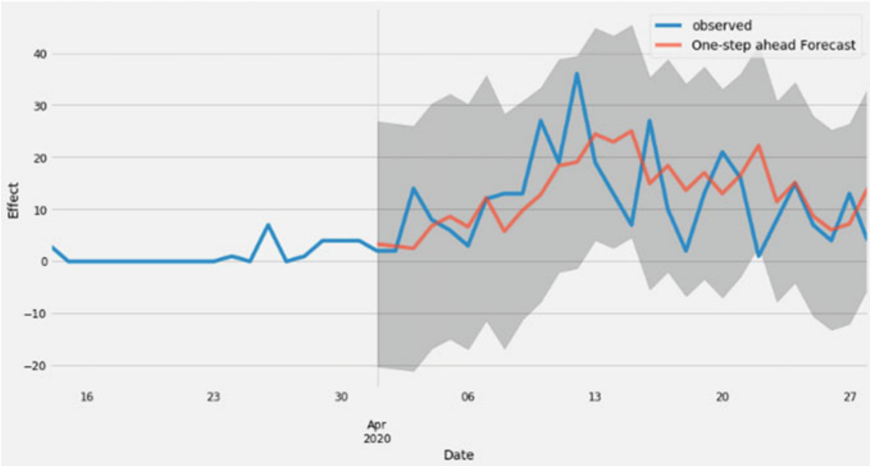



**Fig. 5** Prediction of number of recovered in Kerala state

**Table 3** Predicted date when the number of cases will reduce of various states is displayed

| States of India | Recovery date (Predicted) |
|---|---|
| Andhra Pradesh | 28-07-2020 |
| Bihar | 31-07-2020 |
| Delhi | 19-07-2020 |
| Gujarat | 20-07-2020 |
| Haryana | 21-08-2020 |
| Karnataka | 12-07-2020 |
| Kerala | 27-07-2020 |
| Madhya Pradesh | 30-07-2020 |
| Maharashtra | 21-08-2020 |
| Punjab | 24-08-2020 |
| Rajasthan | 09-08-2020 |
| Tamil Nadu | 31-07-2020 |
| Telangana | 21-07-2020 |
| Uttar Pradesh | 13-09-2020 |
| West Bengal | 18-08-2020 |

had a high RMSE value, we had to consider the accuracy to be 80%. Therefore, from all the values of RMSE, we can estimate that the accuracy of the model can be 90%.

## 6 Conclusion

The effect or impact of the pandemic COVID-19 virus in India was predicted using a statistical model GARCH. The details about the GARCH model was mentioned, and the steps that were involved in the future prediction from the input data are also described. This paper also gives the graphical information about the COVID-19 effect on each state in India with 100 or more cases as per the records till June 15, 2020. As the accuracy of the model was 90%, the predicted information may become true if the rules or precautions which were imposed by the government are strictly followed by the people of respective states, and then the effect may reduce prior to the predicted dates and even the economic condition may come back to normal else there will be a rapid increase in the number of cases and the economy of India will collapse in the near future.

## References

1. Perone G (2020) An arima model to forecast the spread and the final size of COVID-2019 epidemic in Italy. SSRN Electron J. https://doi.org/10.2139/ssrn.3564865
2. Weron R (2014) Electricity price forecasting: a review of the state-of-the-art with a look into the future. Int J Forecast 30. https://doi.org/10.1016/j.ijforecast.2014.08.008
3. Kumar M, Anand M (2014) An application of time series arima forecasting model for predicting sugarcane production in India. Stud Bus Econ 9:81–94
4. Almasarweh M, Alwadi S (2018) ARIMA model in predicting banking stock market data. Mod Appl Sci 12:309. https://doi.org/10.5539/mas.v12n11p309
5. Awajan AM, Ismail MT, Wadi SAL (2018) Improving forecasting accuracy for stock market data using EMD-HW bagging. PLOS One 13(7):e0199582. https://doi.org/10.1371/journal.pone.0199582
6. Adebiyi A, Adewumi A, Ayo Charles (2014) Stock price prediction using the ARIMA model. In: Proceedings—UKSim-AMSS 16th international conference on computer modelling and simulation, UKSim 2014. https://doi.org/10.1109/UKSim.2014.67
7. Armstrong J (2001) Evaluating forecasting. Methods. https://doi.org/10.1007/978-0-306-47630-3_20
8. Chenghao L, Hoi SCH, Zhao P, Sun J (2016) Online ARIMA algorithms for time series prediction. In: Proceedings of the 13th AAAI conference on artificial intelligence (AAAI'16). AAAI Press, pp 1867–1873
9. Taylor S, Letham B (2017) Forecasting at scale. Am Stat 72. https://doi.org/10.1080/00031305.2017.1380080
10. Petrevska B (2017) Predicting tourism demand by A.R.I.M.A. models. Economic Research-Ekonomska Istraživanja. 30:939–950. https://doi.org/10.1080/1331677X.2017.1314822
11. Das S, Ray S, Sen A, Siva, G, Das S (2019) Statistical study on modeling and forecasting of jute production in West Bengal. Int J Curr Microbiol Appl Sci 8:1719–1730. https://doi.org/10.20546/ijcmas.2019.807.204
12. Alsharif MH, Younes MK, Kim J (2019) Time series arima model for prediction of daily and monthly average global solar radiation: the case study of Seoul, South Korea. Symmetry 11:240
13. Totakura V, Devasekhar V, Sake M (2020) Prediction of stock trend for swing trades using long short-term memory neural network model. Int J Sci Technol Res 9:1918–1923
14. Sagar BK, Reddy EM, Ramakrishna S (2018) Network devices data visualization using Weka. Int J Pure Appl Math 118(17):599–608

15. Nugroho A, Simanjuntak BH (2014) ARMA (Autoregressive Moving Average) Model for Prediction of Rainfall in Regency of Semarang-Central Java-Republic of Indonesia
16. Sidiq M (2018) Forecasting rainfall with time series model. IOP Conf Ser: Mater Sci Eng 407:012154. https://doi.org/10.1088/1757-899X/407/1/012154
17. Data Source: https://api.covid19india.org/csv/

# Navigation Assistance and Collision Avoidance for the Visually Impaired "NACVI"

**K. Venkatesh, N. Nalini, M. N. Thippeswamy, Chethan D. Chavan, Sam Jefferey, and Kanitha Tasken**

## 1 Introduction

The WHO estimates that there are about 285 million people in the world who suffer from visual impairments [1]. Being visually impaired constrains the movement of the affected and severely impacts pursuit of their activities. A majority of visually challenged person's use the 'white cane' for avoiding obstacles. The cane is just a simple stick that allows the user to 'feel' obstacles when the stick strikes the obstacle ahead of them. It hasn't been updated since its conception in 1921 [2].The system proposed by the researchers is an amalgamated solution for navigation assistance and collision avoidance for the visually challenged. There have been various approaches with respect to localization and navigation assistance for the visually impaired that uses different sensor and actuators modules. Extensive research has been done to provide a navigational system for the visually impaired. However, there isn't any system that provides a comprehensive solution to the problem. The availability of smart, miniaturized inexpensive sensors in the recent past has made it possible to augment the "**White Cane**" with appropriate obstacle detection sensors like RADAR, LIDAR, Ultrasonic sensors, GPS modules, GSM communication modules, and actuators to detect obstacles reliably even at longer distances which will give sufficient time for

K. Venkatesh (✉) · N. Nalini · M. N. Thippeswamy · C. D. Chavan · S. Jefferey · K. Tasken
Department of Computer Science Engineering, NITTE Meenakshi Institute of Technology, Bengaluru, Karnataka 64, India
e-mail: krishnarao.venkatesh@nmit.ac.in

N. Nalini
e-mail: nalini.n@nmit.ac.in

M. N. Thippeswamy
e-mail: thippeswamy.mn@nmit.ac.in

C. D. Chavan
e-mail: chethan.d.chavan@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_36

the person to avoid the obstacle, in addition in the case of emergencies SOS alert messages can be sent through SMS to registered care givers for appropriate action. An accompanying Smartphone-based application named ***"Kannu"*** has been developed which augments the functions of the White Cane. If the Smartphone is coupled to the White cane then the GPS and GSM modules of the smart phone will be used and the ones on the cane will be redundant (a lower cost version can be configured without these sensors). Kannu provides features like "find cane" could be used for locating the cane if it is misplaced, "Send SOS" alert messages and also aid in navigation. Functions of Kannu are completely voice controlled, for easy operability by the Visually Impaired user.

## 2 Related Work

Recently, quite a few attempts have been made to develop aids to provide collision avoidance and navigation assistance for the visually impaired. A few of these have been discussed below. It is interesting to note the novel ways in which these solutions utilize the different sensors in various modes and configurations to realize the objectives. A few of these solutions have are presented here.

The Navbelt was created by Borenstein and his associates in the University of Michigan [3] as a direction framework which utilizes a versatile robot hindrance shirking framework. The model was developed in 1992 and it comprised of ultrasonic sensors, a PC and a set of stereophonic earphones. The PC gathers data from the eight ultrasonic sensors which generates an ultrasonic beam, and processes the reflections from the obstacles to create a virtual picture by determining the profile of the obstacles (more importantly the outer edges). Based on this picture it guides the user by giving sonic cues using which he can go around the obstacles. The gadget however proved to be very bulky and not user friendly.

Vetteth [4] proposed a gadget was intended for a problem free person on foot route framework. It coordinated a few advances including wearable PCs, picture preparing, sound handling and sound route and running. This gadget centers on realizing a methodology which would permit an outwardly disabled individual to stroll through occupied streets and assist them with distinguishing snags with no difficulty. The gadget utilized a computerized camera to capture the scene that is straight ahead, and this was processed to detect the obstructions and cautions the user through vibratory feedback. The sonar sensors identified hindrances in the user's immediate vicinity. Upon identification, the vibrators would alert him/her with respect to the nearness of the obstacles. The requirement of sufficient light for the cameras to function properly constrains the user from moving out at night and also in dark passages thus limiting the utility of the system.

Zelek's [5] has discussed the design of a wearable device in the form of gloves which has a GPS sensor, Processing element and vibrators embedded into the gloves. The GPS device determines the current location of the user and based on the desired destination to which the user wants to travel the system determines a route. The user is then given appropriate vibratory cues to help him orient himself and move on the determined route. This however does not help in determining the presence of obstacles in the path.

Sharf [6] planned the “Individuals sensor,” which utilizes piezoelectric and ultra-sound sensors so as to differentiate between lifeless objects and those with life (animals, humans etc.). The piezoelectric sensor helps in doing this by detecting the temperature of the object while the ultra-sound sensor will help in determining the distance between the user and the object based on which it warns the user of possible collisions. However, the response time of the sensors and the sensitivity impacts the accuracy of the detection. Crowded places will result in hits frequently and could render the user confused and render him immobile.

The work by Peng et al. [7] presents a monocular approach that utilizes the inbuilt camera on a Smartphone. This framework can recognize obstructions lying on the ground. Tests were conducted on five clients with visual impairment that were supported with vibrations and voice on-request from the Smartphone. The proposed solution required the user to keep the device continuously tilted at an angle of about 45 degrees while he is walking to capture the images of the obstructions lying ahead. This does not allow the hands to be free and in addition has problems with respect to usage in the evenings after it gets dark.

Few other systems based on usage of ultrasonic sensors [3, 6, 8–14] as the main sensing element for obstruction detection. However, all of them suffer from the problem of inaccurate detection during strong gusts of wind, presence of foam and other sound absorbing materials in the surroundings, thus leading to unsafe movement advisories. Most of these also use tactile feedback for giving cues to the user.

Most of these systems address only parts of the problems faced by the visually impaired and hence fail to provide a satisfying solution which does not satisfy all the use-cases. The system discussed in this paper provides an all-weather collision avoidance and navigation system by using the multiple obstacle detection sensors and communication modules. It overcomes the weaknesses of the single sensor configurations outlined earlier and provides a comprehensive and safe solution satisfying all the identified use cases.

## 3 Working Principle

The NACVI smart cane is equipped with RADAR, SONAR and LIDAR sensors to be able to sense objects ahead and also the ditches in the path of the user. The microcontrollers, sensors, actuators, communication modules and power source are embedded into the cane. Sensors such as RADAR, ultrasonic are used to detect obstacle as shown in Fig. 1 and laser is used to detect potholes present in the path. Buzzer and vibrators are used to provide audio and haptic cues to the user. The GPS is used to determine the current location of the user (latitude and longitude), the GSM communication module is used for sending the user’s current location to the

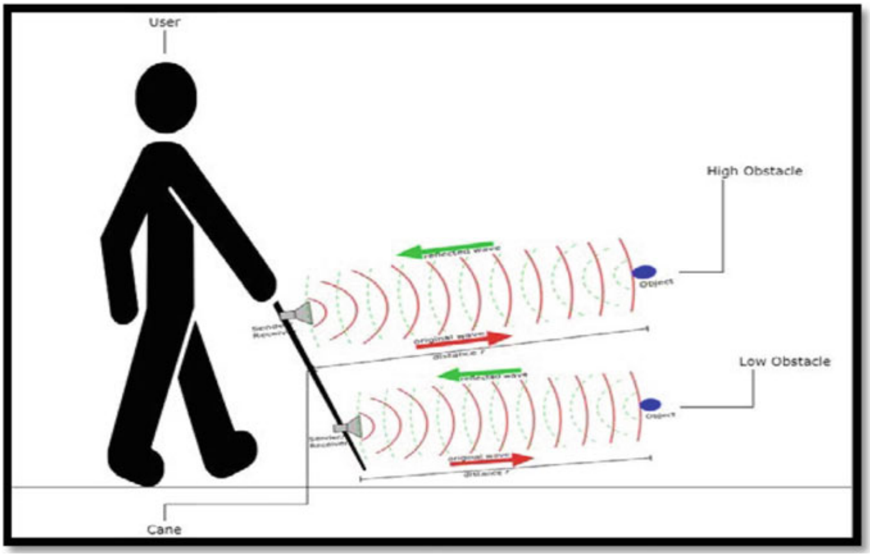



**Fig. 1** Working principle of the obstacle detection sensors

caregiver (buddy) as an alert SMS in case of an emergency, and the Bluetooth module is used to interface the cane to the smart phone (optional). An application.

When an obstacle which could pose a problem to the user is detected by the fusion of the information from all the sensors, a haptic as well as an audio warning will be generated. The alert signal is modulated depending on the distance between the person and the detected obstacle as stated below:

- *Condition 1:* when the obstacle is beyond 250 cm, a low frequency audio and vibratory signal is generated.
- *Condition 2:* when the obstacle is between 100 and 249 cm, a mid-frequency audio and vibratory signal is generated.
- *Condition 3:* when the obstacle is at less than 99 cm, a high-frequency audio and vibratory signal is generated.

## 4 Obstacle Detection, Avoidance, Navigation, Communication, Alert and Feedback

### *4.1 Smart-Cane*

The following are the features present in the designed smart cane: (Fig. 2 shows the physical layout of the main subsystems of the cane.)

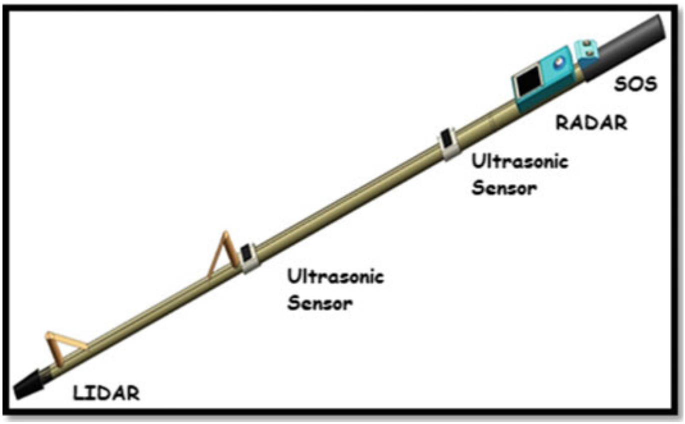



**Fig. 2** Layout of the sensors

### 4.1.1 Obstacle Detection

- A forward-looking ultrasonic sensor transmits an Ultrasonic pulse train and captures the reflected pulses. Based on the TOF (Time of flight) the distance to the reflecting object is determined. This range information is passed onto the fusion module for further processing.
- Forward-looking Frequency Modulated Continuous Wave (FMCW) radar is used to detect multiple obstacles using a carrier 24 GHz RF band signal. A multi-segment TX/RX patch antenna is used to Transmit and receive the Frequency modulated RF signal. The distance to the object is determined by determining the phase difference between the transmitted and received signals in the MTI mode, however if the Stepped Frequency mode of operation is adopted then the distance to the object is determined by the frequency difference between the reflected signal and the one that is currently being transmitted. This processing is done by the module and the range (distance) of the closest obstacle is calculated. This range information is passed onto the fusion module for further processing.

### 4.1.2 Ditch/Pot Hole Detection

A downward looking LIDAR Sensor is used to detect the pothole. This device behaves very much like RADAR except that it beams out a pulsating light beam using a Laser diode; it has an integrated photo-diode receiver which generates a signal when the received signal crosses preset amplitude. The time of flight is calculated based on the time difference between the transmitted signal and the received signal which is then used to determine the distance to the obstacle. This range information is passed onto the fusion module for further processing.

#### 4.1.3 Navigation

- A GPS module is used to capture the cane's current location, providing real time latitude and longitude coordinates enabling real time tracking of the user.
- An RF-ID tag reader is also integrated into the cane. In order to make buildings VI friendly, passive RF-ID tags can be fixed in the corridors and on doors of cabins with unique ids. The RF-tag reader on the cane then can be used for indoor navigation where typically the GPS signals are very weak and one might end up with inaccurate bearings. (Optional feature not shown in Fig. 2).

#### 4.1.4 Communication

The smart cane is provided with a Bluetooth communication support by using the HC04 module. This is used for the information exchange between the smart cane and the Kannu app executing on the associated smart phone.

#### 4.1.5 Alert SOS Button

A button on the cane is used to trigger emergency services. Kannu on receiving this signal acquires the current GPS coordinates [15] and send an alert message to the buddy's phone. Message includes goggle map link containing latitude and longitude of the cane/user to buddy (emergency) contacts saved in Kannu.

#### 4.1.6 Feedback

A Haptic feedback is provided with vibration and Sound to alert the user. Different patterns of vibrations and sounds are used based on the distance between the user and the obstacle.

### *4.2* "Kannu"—*Smartphone App*

#### 4.2.1 Emergency Button

A button on the cane must be long-pressed to trigger emergency services. This includes, sending a message to emergency contacts of the user's location using the Android application.

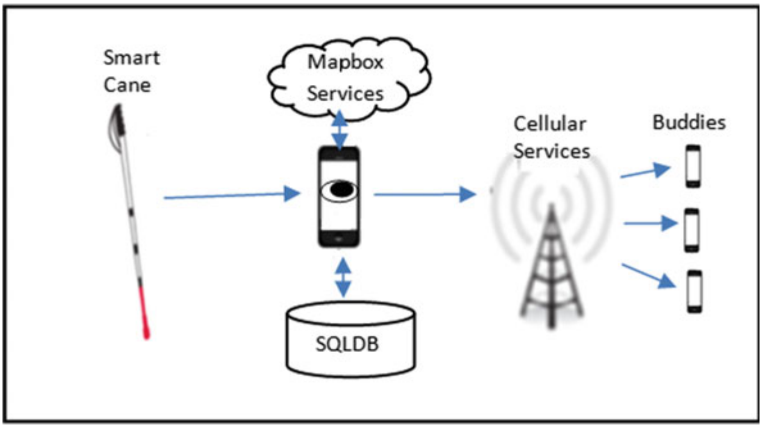



**Fig. 3** Kannu (Eye) system architecture

#### 4.2.2 Find My Cane

Voice command "FIND" to the Android application triggers a beeping pattern on the cane enabling the user to find it using his/her auditory senses as shown in the Fig. 3.

#### 4.2.3 Navigation

Voice command "NAVIGATE" to the Android application, initiates a map interface built on the Map Box SDK. The user speaks the name of the destination and is routed to it using different routing algorithms.

#### 4.2.4 Smartphone

The GSM and GPS modules present in the Smartphone are employed to enable proper functioning of the app and facilitate geocoding and message transfer.

## 5 Components Used for Implementation

### *5.1 Hardware Components*

The block diagram of the Smart Cane and interface of hardware components is as shown in Fig. 4.

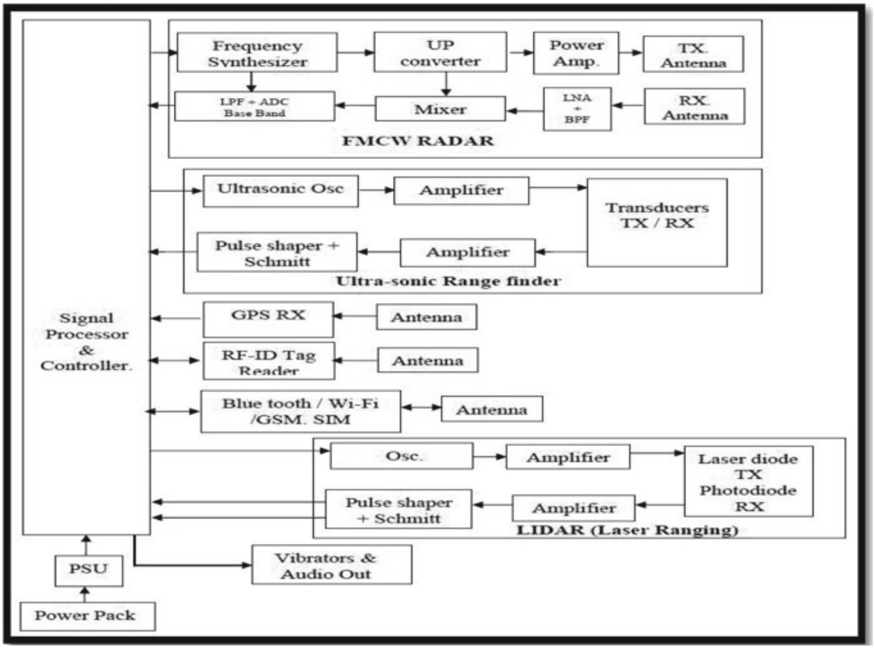



**Fig. 4** Smart cane system block diagram

### 5.1.1 Microcontroller

ATmega328P is a low power CMOS 8-bit microcontroller based on RISC architecture, 32×8 GPR, 32 KB of flash memory, 1 KB EEPROM, 2 KB internal SRAM, 8 bit timer/counter, 10bit on chip ADC, UART, SPI, I2C and 8 MHz internal clock generator.

### 5.1.2 FMCW Radar

BGT24MTR12 is a silicon germanium MMIC for signal generation and reception, operating at 24 GHz VCO, Low noise, and low power consumption 660mW, homodyne quadrature receiver, on chip power and temperature sensors and SPI communication interface.

### 5.1.3 Ultrasonic Sensor

HR-04 generates an ultrasonic sound at a frequency of 40 kHz which transmits through the open air and if there is an object detected it will consider it as an obstacle in its transmission path and it will reflect back to the receiver module, low power device and independent RX TX systems.

### 5.1.4 Laser Sensor

The VL53L0X is a new ASCIC for Time-of-Flight (ToF) used in laser-ranging module totally invisible to human eye. Measuring a range up to 2 m, 940 nm laser VCSEL driver, I2C interface for communication programmable address, size of 4.4 $\times$ 2.4 $\times$ 1.00 mm and low power consumption device.

### 5.1.5 GPS

NEO-6 module is a standalone GPS receiver its miniature of 16 $\times$ 12.2x2.4 mm package which is compact and compatible for our project. The 50-station U-Blox 6 positioning engine boasts a Time-To-First-Fix (TTFF) of less than 1 s empowering it to discover satellites right away. Inventive structure and innovation stifles sticking sources and mitigates multipath impacts, giving NEO-6 GPS collectors fantastic route execution even in the most testing conditions.

### 5.1.6 GSM

SIM800C is a quad-band GSM/GPRS module working at 850 MHz and 900 MHz, 1800 MHz and 1900 MHz. dimension is of 17.6 $\times$ 15.7 $\times$ 2.3 mm SMT package which is compact and compatible foe our project. UART protocol, 1 SIM card interface and power saving technique sleep mode at 0.6 mA power consumption.

### 5.1.7 Bluetooth Communication Module

HC-05 is a low power module which is a full-duplex wireless functionality in this project. Range is of $<$ 100 m, UART communication, uses FHSS and supports multiple baud rate.

### 5.1.8 RF-Id Tag Reader

An EM-18 is an active short range, low power RF-Id reader has been integrated into the system. IZOKEE reader with compatible tags is used. The reader module sends the information to the processor on a serial communication channel.

#### 5.1.9 Switch

Push button switch with quick denounce technique.

#### 5.1.10 Buzzer

Low power transducer buzzer.

#### 5.1.11 Vibration Motor

Low power DC motor.

### *5.2 Software Components*

The Signal Processing Chain of the system is as shown in Fig. 5.

#### 5.2.1 FMCW

The controller triggers the ramp sweep generator of the Digital Synthesizer which will generate CW signals with monotonically increasing frequency based on the

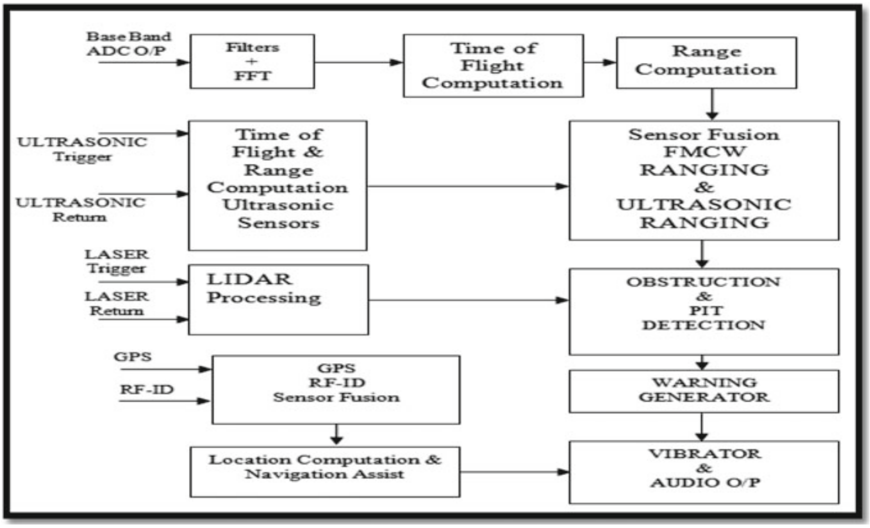



**Fig. 5** Signal processing chain

preset slope covering the whole frequency range. The received digitized base band (post mixing) is a measure of the phase shift between the transmitted & reflected signal. A 1024/2048 point FFT is performed on this signal. The frequency of the received signal is directly proportional to the range of the object.

### 5.2.2 Ultrasonic

If the Ultrasonic sensors are used for the short distance measurement then this module will generate the trigger for generating the burst of ultrasonic signals. The receiver will be kept enabled. The time between the trigger and the received signal is determined. Based on this the distance between the transmitter and the target is computed.

### 5.2.3 Pit Sensing

The time of flight between the transmitted Laser signal and the Received signal is determined in this module. This TOF reflects the distance to the object reflecting the narrow beam signal. Sudden increase in the depth (as perceived by the LIDAR) is an indication of either a pit or stairs and hence will be flagged by this module.

### 5.2.4 Sensor Fusion

The range information of various obstructions and pits as detected by the Radar sensors, Ultrasonic sensors, and LIDAR will be combined to form a dynamic picture of the surroundings. This combined picture is interpreted to determine if any of the detected obstructions will pose danger to the person. If so, the appropriate warnings and cues will be generated.

The Location information will be acquired from the GPS unit and the RF-ID unit. Consistent and accurate Location information is created by combining the information from these various sources. This is regularly updated by this module.

The Navigation assistance module will give appropriate guidance to the person based on the present location and the destination.

### 5.2.5 Emergency Message

The location information is transferred in case of an emergency through the selected communication links: Bluetooth, Wi-Fi, or Cell phone.

# 6 Physical Packaging Design Prototype

The layout of the Radar: BGT24MTR12 module is shown in Fig. 6. This has been also integrated into the smart cane as a separate module.

The prototype was designed using individual components as mentioned above. Figure 7 Shows one configuration of the prototype.

## 6.1 *Software Development Environment*

### 6.1.1 Arduino-IDE

The Arduino Integrated Development Environment (IDE) is a cross-stage application. The programs for the Arduino-AT-Mega 328P processor used in all the processing modules can be developed using this Arduino IDE. It also supports 3rd party libraries and customization of the same. The environment also supports the upload function to transfer the code to the Arduino development board/AT-Mega processor as required and also to test the same.

### 6.1.2 Android Studio

Android Studio 3.0 is an IDE for android app development it has a built-in code editor and emulator to virtually visualize and emulate the app developed inside the tool environment itself. Subsequently, it can be ported onto the Smartphone environment. Most of the app logic testing can be done using the simulator. The app is then uploaded to the target Smartphone in order to test all the sensor related

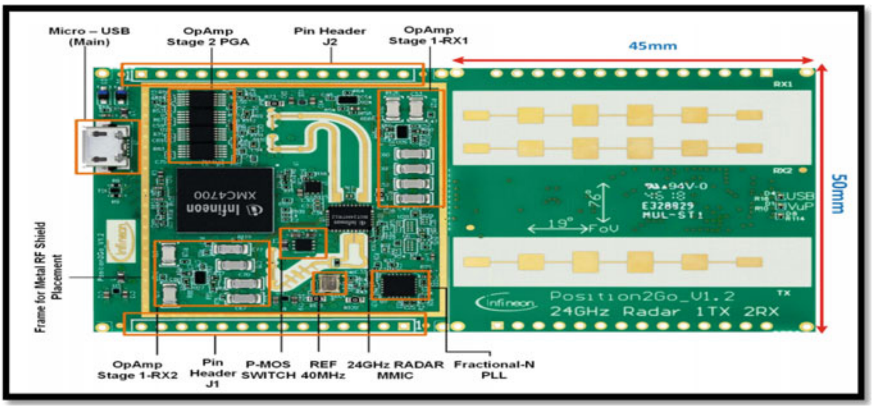



**Fig. 6** 24 GHz radar development board

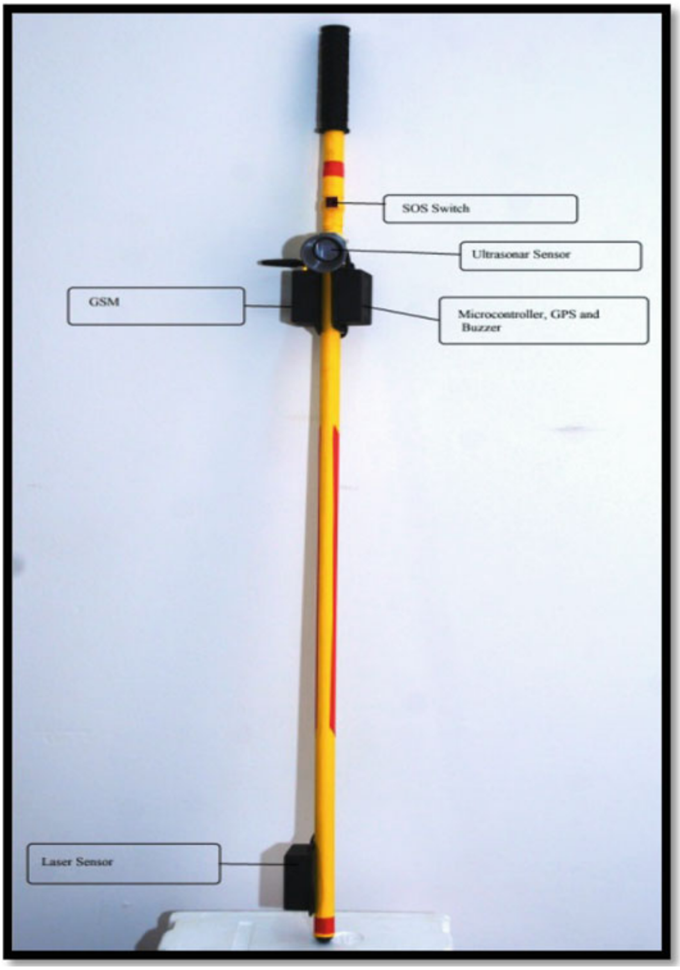



**Fig. 7** Ultrasonic version of the smart-cane

I/O operations and complete the integrated testing of the application. Since, in the application being discussed voice operated UI is being used it might be necessary to try out the operations using different voice patterns.

### 6.1.3 Infineon Toolbox

The RADAR signal processing chain is built using the Infineon Toolbox IDE for the BGT24MTR11 module. This helps in visualizing the obstacle maps and customizing the Signal processing software. The same is shown in Fig. 8.


**Fig. 8** Radar Obstacle map visualizer and SW development environment

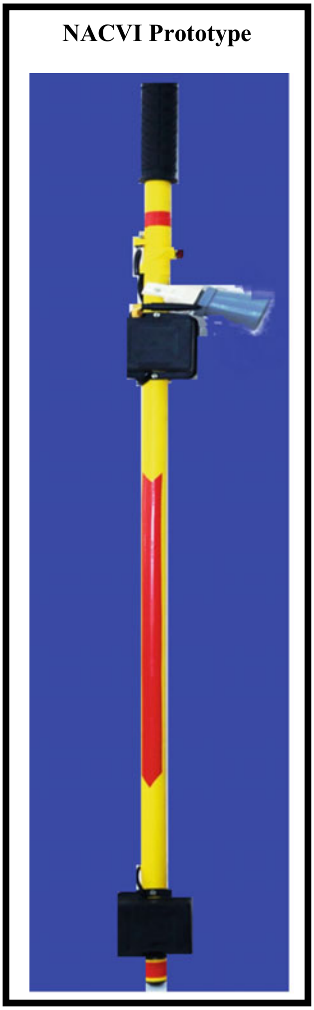
## 7 Result

The prototype was tested for specific indoor and outdoor use cases and was found to be working satisfactory. User testing of the cane was also conducted in a school for the blind. Pictures of the same are shown below in Fig. 9. Prototype Cane which was used for testing and the users found it very useful and have also given suggestions which will be reviewed and incorporated in the next version of the device.

## 8 Conclusion

Designing a device that will solve all the problems associate with navigation and collision avoidance for the visually impaired is not a simple task. The complex use cases, environmental and affordability issues all play an important part in choosing the sensors. The aforementioned design is an attempt at building a customizable and affordable solution for the same. The initial trials have been very encouraging but needs to be followed up with extensive trials. It should be noted that cost, weight and usability were the key driving factors for choice of the sensors and processing modules.

Fig. 9 Prototype




## 9 Future Work

Extensive environmental and user testing along with a production house is planned to be performed so as to derive the requirements for a mass production version of the device. The production unit will also take into consideration the fact that different functional configurations may have to be developed to meet certain price-performance criteria. The signal processing software that is being currently used can be augmented to learn from the specific usage pattern of the particular user and the environment in which it is being used to provide more accurate cues to the user. A self-learning module can be incorporated at a later stage. Additionally, a Built-in-test module will be added which can test all the modules of the cane every time it is switched on and inform the user about the health of the cane, this will build the confidence to the user about its safety and usability.

**Acknowledgements** Project funded by Science for Empowerment and Development Division, DST, MHRD, Government of India. **Sanctioned No.: SEED/TIDE/038/2015/G**.

## References

1. World Health Organization (2017), Vision impairments and blindness fact sheet, October 2017. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs282/en/. Accessed January 2018
2. Strong P (2009) The history of the white cane. Tennessee council of the blind, January 2009. [Online]. Available: http://www.acb.org/tennessee/white_cane_history.html. Accessed January 2018
3. Shoval JBAYKS (1998) Mobile robot obstacle avoidance in a computerized travel aid for the blind. IEEE Trans Biomed Eng 45(11):1376–1386
4. Vetteth PGKDKG (2013) Collision avoidance device for visually impaired. Int J Sci Technol Res 2:185–188
5. Zelek MFJS (2013) Wearable tactile navigation system. United States of America Patent US20130218456 A1, August 2013
6. Sharf SRAJ (1998) The people sensor: a mobility aid for the visually impaired. In: Proceedings of the 1998 digest of papers 2nd international symposium on wearable computers, 1998
7. Chandrasekaran SHSMDMQPCK-GN (2017) Visual impairment predicts poor physical functioning among middle-aged women. Innov Aging 1(1):1124
8. Milios BKAKASSE (2003) Sonification of range information for 3-D space perception. IEEE Trans Neural Syst Rehabilitation Eng 11(4):416–421
9. Magatani KSAKYK (2001) Development of the navigation system for the visually impaired by using optical beacons. In: Proceedings of the 23rd annual international conference of the IEEE engineering in medicine and biology society, 2001
10. Meijer PBL (1992) An experimental system for auditory image representations. IEEE Trans Biomed Eng 39(2):112–121
11. Kavraki NGBAD (2001) An intelligent assistant for navigation of visually impaired people. In: Proceedings of the 2001 IEEE 2nd international symposium on bioinformatics and bioengineering conference, 2001
12. Sainarayanan RNASYG (2007) Fuzzy image processing scheme for autonomous navigation of human blind. Appl Soft Comput J 7(1):257–264
13. Fang JXF-lBAL-HZZ-G (2007) Audio man: design and implementation of environmental information data mapping. Chin J Ergon 2
14. Nie JRZLM (2009) Sound view: an auditory guidance system based on environment understanding for the visually impaired people. In: Proceedings of the 31st annual international conference of the ieee engineering in medicine and biology society: engineering the future of biomedicine (EMBC '09), 2009
15. Android Developers (2018) Location strategies, [Online]

# Health Review and Analysis Using Data Science

**Debashish Dutta, Shivarpan Das, Aritra Nath, Abhyuday Kaushik, and P. Shobha**

## 1 Introduction

Hospitals are trying to adopt state of art technologies into their sphere for technological advancements and cost reductions in medical equipment. They are also trying different measures to diagnosis and predict the future outcomes of diseases. The need of good, interactive and intelligent system for knowledge discovery is rapidly growing. The accuracy and speed are essential for any intelligent system to work and bring useful results. This will help in classifying the disease accurately [1]. The traditional practices or approaches for classifying the disease are not suitable for the current data size since its increasing in a rapid pace [2]. The accuracy and effectiveness of traditional classification algorithms fail when diseases with the characteristics of multiple similar treatment stages, various symptoms, and multi-pathogenesis are fed to the algorithms. Therefore, suitable prediction models are required to accurately classify the disease symptoms. Also, the healthcare industry generates terabytes of data every year. The medical documents contain a lot of data regarding the patients. For extracting and analyzing this amass volume of data is tricky and complex [3]. This data may contain the patient's history, past medical information and much more. To deduce and analyze this data, intelligent systems plays a major role for not only in an interactive way but helps it keep organised which may be difficult for normal human being. The disease prediction and analyzing also plays a major role in healthcare since it can help in deducing the large amount of data and give accurate predictions.

The health disease prediction and analysis were previously done using past records by humans but as the data size started to increase it was difficult to track records and accurately classify a particular disease due to error or mishandling of data.

D. Dutta · S. Das · A. Nath · A. Kaushik · P. Shobha (✉)
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India
e-mail: shobha.p@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_37

But as technological advancement arrived in this sector, there has huge boost in collecting and analyzing the data. Previously traditional or naive algorithms were used to classify data but as data started to grow larger they started to show low performance. The traditional algorithms were also failing when diseases with the characteristics of multiple similar treatment stages, various symptoms, and multi-pathogenesis are fed to the algorithms. Thus improving upon the speed and accuracy were essential to accurately classify a disease.

## 2 Literature Survey

It is an open fact that in recent years the health care industry have done tremendous headway may it be in the field of disease detection or the arena of its cure, may it be gathering of patient information to processing of it to get meaningful and deeper insight we have come a long way [4]. The medical industry produces humongous amount of data and filtering out the useful information from patient's record is indeed a tedious but knotty work which actually needs both technical and analytical excellence. In this case lets take the patients symptoms into consideration. Many systems have been but which takes the patient symptoms as input and try to detect the disease he/she may be suffering from. But, most of these are disease specific and lack accuracy which in a serious matter of concern.

Let's consider few examples where previously some remarkable work related to health prediction has been done earlier. The main aim of these models was to process this bulky, voluminous data and predict disease in order to protect people [5, 6].

One of the most renowned is the health care support for Swine flu it extensively uses the Naive Bias algorithm and categorized the patients into least possible, probable and most probable classes. The major issue with this support system was the data used for categorization was very limited in count.

Another interesting system is a collaborative filtering technique which is based on user's feedback to produce items description. Here the user reviews are analyzed and processed using Core NLP and there closeness is computed with the help of K-NN.

Following it is a Naive Bayes classification approach for dermatological conditions. It processes the patient's attributes and provides the possibility of eight potential skin related diseases. It can also be used to predict other diseases but gives very low accuracy. Moreover, it does not come with any recommendation system or suggestion provider [7].

## 3 Technology Used

Front End Tools:

*HTML:* Hypertext Mark-up Language (HTML) is the standard mark-up language for documents that we view on the web. The HTML gives structure to a web page.

It is assisted with CSS and javascript for givig the web page a creative design. The latest version of HTML5 even supports built-in error checking mechanism that could only be accomplished via JavaScript previously.

*CSS*: Cascading Style Sheets (CSS) is a style sheet language used for describing the look and appearance of a document written in a mark-up language like HTML. Without the CSS, the HTML code looks just give the information. CSS gives a attractive design to the web page which keeps the reader engaged.

It can also enable multiple web pages to share the same formatting by specifying the relevant CSS in a separate file and reduce complexity and repetition in the structural content.

*JavaScript:* JavaScript is a scripting or programming language that allows you to implement complex features on web pages. It is the third layer of the layer cake of standard web technologies, two of which (HTML and CSS). Alongside HTML and CSS, JavaScript is one of the core technologies of the World Wide Web (WWW). JavaScript renders web pages in an interactive and dynamic fashion. This allowing the pages to react to events, exhibit special effects, accept variable text, validate data, create cookies, detect a user's browser, etc.

Backend Tools:

*SQL Alchemy*: Flask SQLAlchemy is an ORM tool which establishes the relationship between the objects and the tables of the relational databases. SQL Alchemy provides a full suite of well-known enterprise-level persistence patterns. It is designed for efficient and high-performing database access. The object-relational mapping is the technique of storing python objects into the database tables without writing the raw SQL queries.

*Flask*: Flask is a web application framework written in Python that provides a web application developer to write applications without having to bother about low-level details. Flask is basically based on the Werkzeug WSGI toolkit and Jinja2 template engine. Both belong to Pocco projects.

*Google Text-To-Speech (GTSS)*: gTTS expanded as Google Text-to-Speech is a Python library and CLI tool to interface with Google Translate's text-to-speech API. It is further written down in an mp3 file which is used in voice analysis and manipulation [8]. It simply pre-generate Google Translate TTS request URLs to feed to an external program.

*SKLearn*: Scikit-learn which used to be formerly called 'scikits learn' and is also known as sklearn is a machine learning library in python to interpret, analysis adn predict data. It provides feature of various classification, regression and clustering algorithms including random forests, support vector machines, k-means, gradient boosting.

## 4 Design

The proposed model first takes the required symptoms from the user both using the option interface and voice enabled feature as required. After detecting the symptoms, it gets transmitted to the machine learning model which evaluates the particular outcome from the symptoms and gives the data to the web page. After receiving the predicted output, we parse the description from a particular site and showcase it. This model is very useful for someone who is in need of knowing the common ailments they are suffering from [9]. Also the voice feature makes the work easier for the users since they only have to speak the symptoms to the device directly and prediction will be displayed on the page. The proposed system will not only predict the disease that a person may suffer from but comes with multitude added features as shown in the flowchart. It will provide suggested balanced diet, exercise and healthy lifestyle information. Moreover, taking the current pandemic situation as of in the year this paper was written, under consideration an purely dynamic covid19 tracker has been added which shows the state wise count of the confirmed, active, recovered and deceased cases across the nation. The system also displays the potential causes and precautions related to covid19 disease (Fig. 1).

## 5 Implementation

Health Review and Analysis Using Data Science consists of primarily two things namely prediction of disease and precision analysis that is calculating the percentage of occurrence of that particular disease. The prediction part consists of three primary steps: symptoms fetching, narrow down symptoms and disease prediction. Firstly, the patient needs to register to our online portal where after entering the credentials he can sign into our web app using the patient mode login. Once, the patient logs in he is given an option of entering all the symptoms like breathlessness, high blood pressure, headache, stomach-ache etc. in a suggestive drop-down menu, also the patient can speak the symptoms using our voice feature button. After getting the symptoms it narrows down disease and predicts the disease using the data fetched in the previous process and uses the best disease prediction model to give us the name of the disease the patient may suffer from.

Here we are using five machine learning algorithms:

i. Decision Tree
ii. Naïve Bayes
iii. SVM
iv. Logistic Regression
v. Random Forest

As we predict the disease, we also give the patient the required information regarding it and redirect them for the preventive measures and further information of the disease to a recognized medical site.

According, to the second major objective which is the precision analysis, after rigorous analysis Support Vector Machine (SVM) was chosen as the prime algorithm based on performance and accuracy. The predicted value of SVM that is our prime

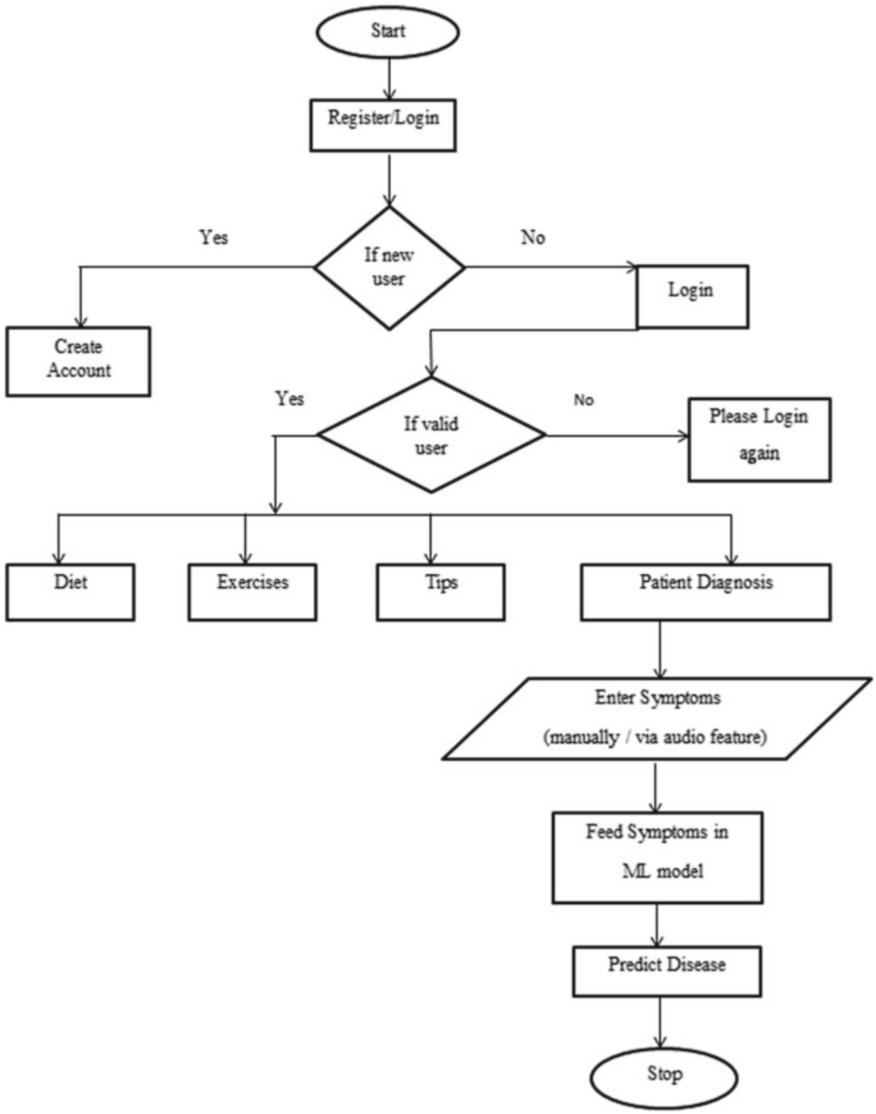



**Fig. 1** Flow chart

algorithm is being compared with the predicted values of other algorithms, following this we take the average of all the like and alike values. Hence, display it as precision.

Also, they can navigate through the diet, exercises and healthy lifestyle tips for proper care and management in their daily routine. As in the recent times, Covid-19 created a huge pandemic and is a major concern across the globe, the app has a tracker particularly for Indian states which gets updated every day since the data is gathered from a official government site which makes it purely dynamic in nature. With that it gives the potential causes of this detrimental disease and recommended tips to take the necessary precautions.

## 6 Results and Discussions

In the present context and considering the current scope of this project, we addressed the primary goal of our project that was to provide remote patients a viable option to understand the disease he/she might be suffering from. These goals were fulfilled by making use of a web application. In the future we are looking at two possible improvements to our product. These being the following –It would be interesting to build a mobile application for our product apart from the web application. The mobile application for the same could be distributed to prospective customers via online app stores [10].

Soon we plan to distribute our application in medical institutions with the intention that when a user gets a complete report of the ailment, it would also be possible to schedule an appointment with the concerned doctor of the medical institution directly from our application without any further interaction. Therefore, it would make our application's usefulness to be two-fold where a user could use this application to predict his ailment as well as share the complete report with the concerned doctor whilst booking an appointment seamlessly all from the same application (Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11).

## 7 Conclusion

The web application proposes a remote way of medical diagnosis. The evolution of data is growing rapidly so that we can learn from it and apply significant patterns to it to detect the cause and take the necessary precaution at an early stage. The way we used big data, machine learning and cloud computing in the healthcare sector will bring a good impact in terms of health and lifestyle in society. We tend to forget the lifestyle required for a healthy being but with this we tend to bring it closer to people as we also give the necessary recommendation for good health. It offers an overview of a patient's health, permitting a more accurate diagnosis and improved patient care.

**Fig. 2** shows the patient registration page

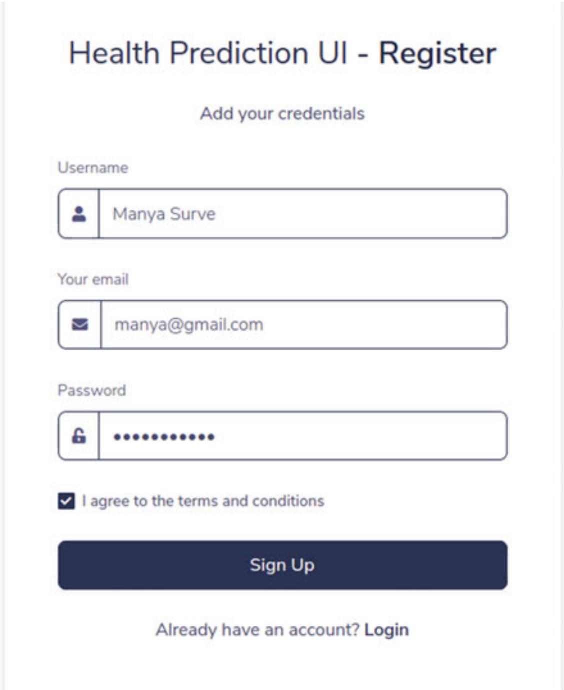



**Fig. 3** shows patient login form page

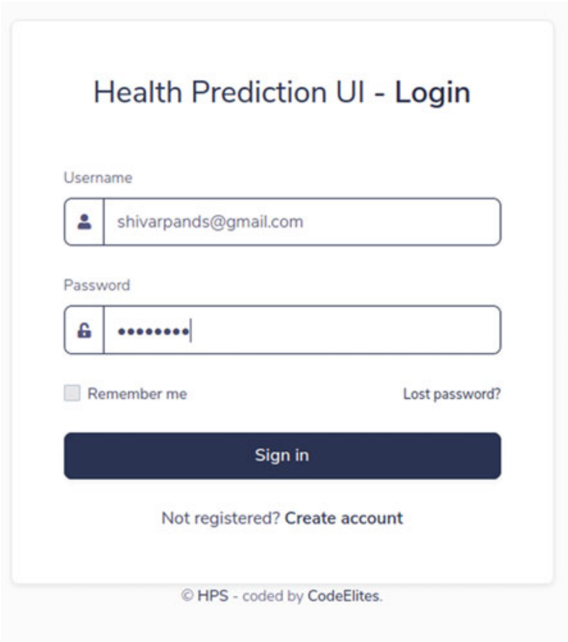



**Fig. 4** shows the Input the user has to provide either from the drop-down menu or using the audio based feature

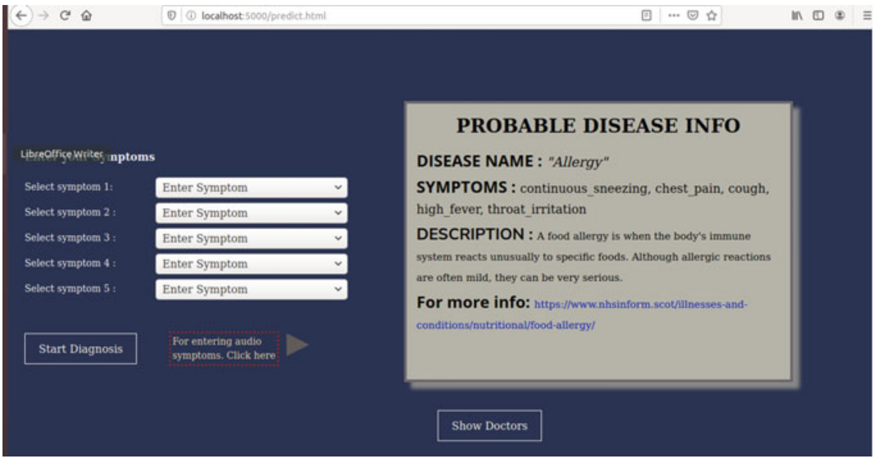



**Fig. 5** shows the output which the probable disease caused by the symptoms provided


**Fig. 6** depicts the symptoms being input through patient's voice

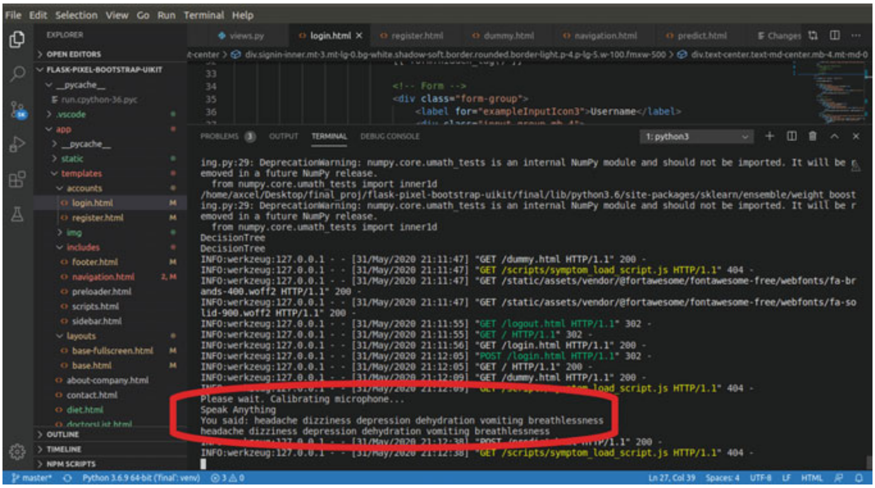



**Fig. 7** shows the voice input shown in backend

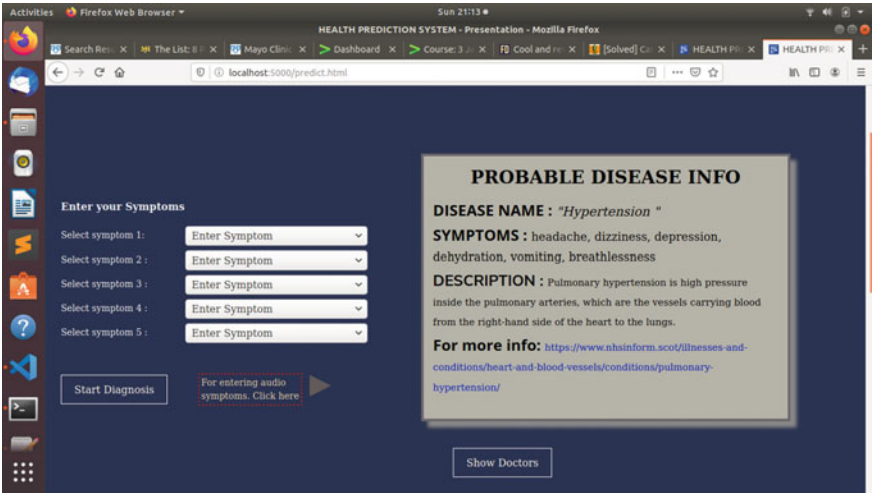



**Fig. 8** depicts the predicted and probable disease output

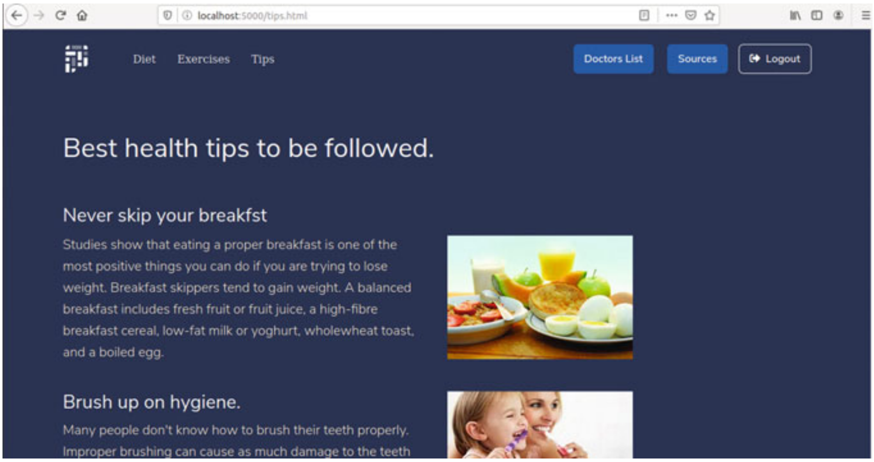



**Fig. 9** shows the suggested health tips

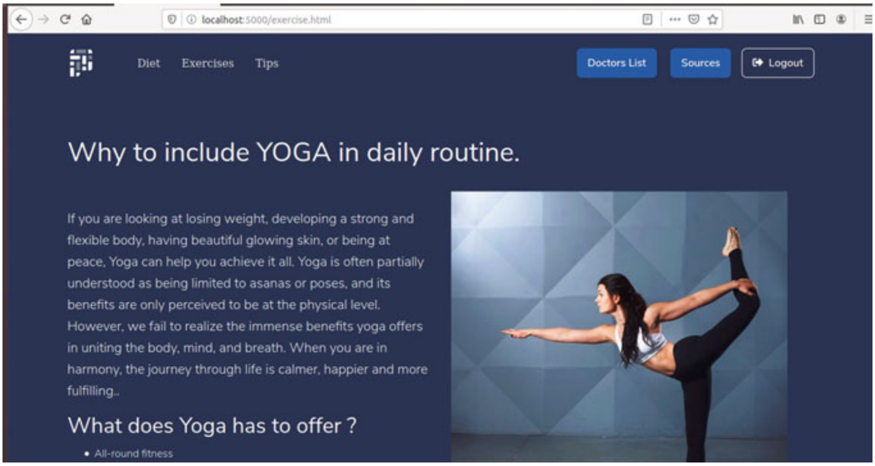



**Fig. 10** shows the suggested exercises

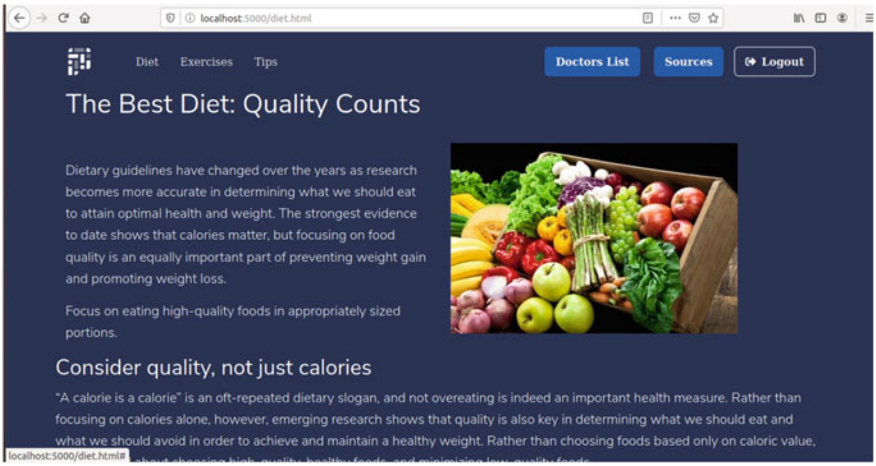



**Fig. 11** shows the suggested diet tips

# References

1. Zheng L, Zhu F, Huang S, Xie J (2017) Context neighbour recommender: integrating contexts via neighbours for recommendations. Inf Sci 414:1–18
2. Chen J, Li K, Tang Z, Bilal K, Li K (2016) A parallel patient treatment time prediction algorithm and its applications in hospital queuing-recommendation in a big data environment. IEEE Access 4:1767–1783
3. Bilal AK, Zhang L, Khan SU (2015) A cloud based health insurance plan recommendation system: a user centred approach. Future Gener Comput Syst 43–44:99–109

4. Gujar D, Biyani R (2019) Disease prediction and doctor recommendation system. Retrieved Int Res J Eng Technol (IRJET)
5. Huang F, Wang S, Chan CC (2012) Predicting disease by using data mining based on healthcare information system. IEEE
6. Isinkaye FO, Fola Jimi YO, Ojokoh BA (2015) Recommendation systems: principles, methods and evaluation. Elsevier 261–273
7. Chen J, Li K, Rong H, Yang N, Bilal K, Li K (2018) A disease diagnosis and treatment recommendation system based on bigdata mining and cloud computing. Accessed at https://arxiv.org/pdf/1810.07762.pdf
8. GTTS, What is Google Text to Speech? https://gtts.readthedocs.io/en/latest/
9. Pandey SC (2016) Data Mining techniques for medical data: a review. IEEE
10. Yao Q, Tian Y, Li P, Tian L, Qian Y, Li J (2015) Design and development of a medical big data processing system based on Hadoop. J Med Syst 39(3):1–11

# Efficiently Revocable Identity-Based Broadcast Encryption Using Integer Matrices as Keys

**B. S. Sahana Raj and V. Sridhar**

## 1 Introduction

Identity-based encryption [1] was first introduced by Adi Shamir in 1984. In identity-based encryption (IBE), the public key of a user is derived from his/her unique identity, say email address, phone number, etc. This avoids the need for a public key certification authority. On the other hand, the IBE system requires a private key generator (PKG) that generates the matching secret keys corresponding to the user's public identities. In IBE, only those receivers who have the secret decryption keys can decrypt the encrypted message.

Broadcast encryption (BE) was introduced by Fiat and Naor [2] in 1994. It is basically a one to many secure data transmission. In BE, a sender encrypts a common message and broadcasts the ciphertext to the registered users who in turn decrypt it using their respective private keys. A salient feature of BE is the provision for revocation of selective users based on the agreed terms and conditions.

Identity-Based Broadcast Encryption (IBBE) is the combination IBE and BE. Delerablee [3] proposed the first Identity-Based Broadcast Encryption (IBBE) in 2007 using the bilinear mapping, Here, a message can be encrypted such that any specified subset of users can decrypt it.

B. S. Sahana Raj (✉) · V. Sridhar
Department of Electronics and Communication Engineering, PET Research Foundation, PES College of Engineering, Mandya, India

Nitte Meenakshi Institute of Technology, Bengaluru, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_38

### 1.1 Revocation Methods

In IBBE, revocation of one or more specific users is a necessary requirement to take care of the expiry of time bound contracts or when the corresponding secret keys have been compromised by adversaries or for any administrative reasons. Revocation can be direct or indirect [4]. In direct method, the sender has the revocation list, and the encryption is carried out to exclude the revoked members. Here, the secret keys of the users are not changed. In indirect method, the keys of the users are periodically updated so as to disable the keys of the revoked members. A traditional indirect method of revocation is described in [5]. Yet another scheme is to use trusted security mediators to block the decryption process of revoked users [6, 7]. These existing methods incur substantially higher processing and communication overhead. To overcome this disadvantage, we propose a scheme, where user revocation information is embedded in the ciphertext itself. Thus, it is direct revocation.

### 1.2 Contributions

Present study proposes a new technique to realize IBBE using matrix keys designated as IBBE-MK. It has efficient revocation method. In IBBE-MK, integer vectors and matrices in the finite field $Z_p$ are used as the secret keys. The ciphertext is derived from the plaintext and the master secret key generated by the PKG such that only the intended authorized (unrevoked) users can decrypt the ciphertext.

## 2 Related Work

In [3], the author has described the first IBBE scheme where the ciphertext has constant size. This method uses key encapsulation mechanism where the broadcast ciphertext encrypts a short symmetric key which in turn is used to encrypt long text messages. In [8], constant-sized ciphertext encryption with efficient revocation is presented. The authors use the functional encryption schemes and "n-equation technique" to achieve their objective. In [9], the author has used IBBE scheme based on external Diffie–Hellman assumption with ciphertext and key sizes sub-linear with respect to the number of receivers. In [10], collision resistant IBBE is described that provides constant-sized ciphertext and private keys. In [11], receiver privacy preserving IBBE is presented. Here, the identities of the receivers are protected from malicious attackers. In [12], Anonymous IBBE with revocation for file sharing is described. In [13], Boneh et al. have described collision resistant broadcast encryption schemes for stateless receivers. Here the ciphertext and key sizes are of order $O(\sqrt{n})$ where $n$ is the total number of receivers. In [14], the authors have used dual-mode broadcast encryption which provides flexibility in the process of revocation. In [15], the authors have presented a new method for IBBE with constant ciphertext size. The proposed method provides protection against decryption key exposure. In [16], generalized IBBE is implemented using bilinear Diffie–Hellman exponents. In [17], IBBE is realized using proxy re-encryption. In [18], hierarchical IBBE is described with efficient revocation. In this case, the receivers form a tree topology, and the revocation of child nodes are delegated to the parent nodes.

In [19], recipient revocable IBBE is presented where the revocation is implemented by a broadcaster which has received the encrypted content without decryption. In all the above cases, [3–19], encryption and revocation are implemented based on bilinear pairings which are computationally expensive and have a high level of communication overhead. To overcome this, we use matrix-based method to implement IBBE. At present, a few matrix-based IBBE schemes are available in literature. These methods are based on learning with errors (LWE) in lattice cryptography.

Agrawal and Boyen [20] have presented IBBE in the standard model using hard lattices based on the pre-image samplable functions. In [21], Gentry, Halevi and Vaikuntanathan (GHV) have presented IBE encryption process that uses GPV {Gentry, Peikert and Vaikuntanathan [22]} trapdoor function. GHV method achieves shorter ciphertext expansion ratio. In [23], an IBBE is presented with a short public key. This is achieved by encoding the user IDs as vectors instead of matrices as in [20]. In [24], the authors have described adaptively secure IBBE using lattices. In [25], Forward Secure Identity-Based Broadcast Encryption Scheme using integer Lattices. The Forward-Security provides protection to past private keys of a user when the current private key of that user is compromised.

## 3 Symbols, Notations and Definitions

The proposed Identity-Based Broadcast Encryption using Matrix Keys (IBBE-MK) scheme uses integer vectors and matrices in $Z_p$ as the secret keys.

### *3.1 Users and Their Identities*

Let the number of users who are the receivers of the data be $N$. During initialization, the individual users are enumerated as $\text{User}_1$, $\text{User}_2$,…,$\text{User}_N$ in some order, and the same order is maintained throughout. For convenience, these users are designated by the short symbols $u_1$, $u_2$,…,$u_N$. The symbol $\text{u}_j$ acts as the symbolic name of the $j$th user. The set of all users is as represented by$U$ as,

$$U = [u_1, u_2, \ldots, u_N] \tag{1}$$

The distinct IDs of the users (say email addresses, biometric information or any suitable unique identifier) are strings denoted by $\text{ID}_1$, $\text{ID}_2$,…, $\text{ID}_N$. The length of

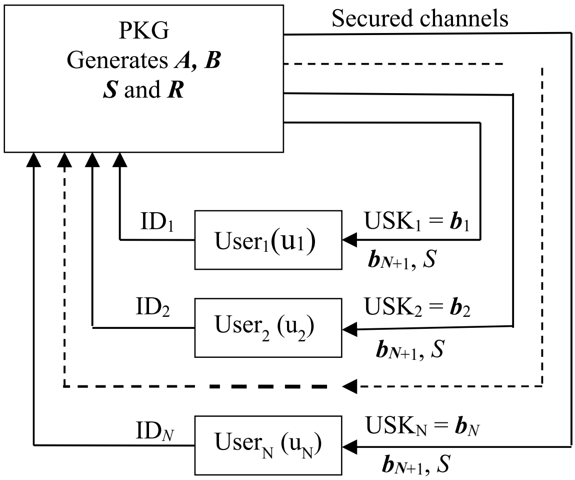



**Fig. 1** Basic layout of PKG and users in IBBE-MK

IDs need not be same. In our scheme, the public keys of the users are their own IDs. Thus, the User Public Key (UPK) set is

$$\text{UPK} = [\text{ID}_1,\ \text{ID}_2, \ldots,\ \text{ID}_N] \tag{2}$$

## *3.2 Private Key Generator*

The function of the trusted PKG is to create the private (secret) keys of the users based on their public keys and the master secret key of PKG. The basic layout of the PKG is shown in Fig. 1.

## *3.3 Registration*

Initially, all the users have to register themselves at PKG. Each user submits his/her public key (ID) and the chosen {user-name, password} pair for user authentication. When the public keys arrive, the PKG collects the public keys and other details and stores them in the order of registration in a table designated as user information register (UIR) whose format is shown in Table 1.

Additional entries in the last two rows are used for signature insertion and randomization as will be explained later. The last column with heading RL represents the revocation list which is addressed in further section.

**Table 1** User information register (UIR)

| Sl. no | UPK (IDs) | Symbolic name | Username | Pwd | USK (**B**) | RL (**E**) |
|---|---|---|---|---|---|---|
| 1 | $ID_1$ | $u_1$ | – | – | $\boldsymbol{b}_1$ | $e_1$ |
| 2 | $ID_2$ | $u_2$ | – | – | $\boldsymbol{b}_2$ | $e_2$ |
| – | – | – | – | – | – | – |
| $N$ | $ID_N$ | $u_N$ | – | – | $\boldsymbol{b}_N$ | $e_N$ |
| $N+1$ | – | – | – | – | $\boldsymbol{b}_{N+1}$ | – |
| $N+2$ | – | – | – | – | $\boldsymbol{b}_{N+2}$ | – |

### *3.4 Setup*

Let the security parameter be $k$ (length of the elements of the key matrix in bits). The PKG chooses a prime number $p$ whose size is $k$ bits. This $p$ forms modulus of the finite field$Z_p$. We use modular arithmetic/algebra in all our calculations with modulus $p$ which is relatively a large prime number. All our variables and parameters belong to the prime field $Z_p$. Thus, all the numbers used for encryption and decryption belong to the integer set $\{0, 1,\ldots,p-1\}$. Parameter $p$ is made public and available to all users.

### *3.5 Key Generation*

The PKG generates the master secret key (MSK) and the user secret key (USK) matrix. In proposed scheme, there is no public key.

**Generation of MSK**: The MSK is a random integer non-singular matrix in$Z_p$ represented by symbol $\boldsymbol{A}$. The size of $\boldsymbol{A}$ is $(N + 2) \times (N + 2)$. That is $\boldsymbol{A} \in Z_p^{(N+2)\times(N+2)}$.

**Calculation of USK**: The USK matrix is represented by symbol $\boldsymbol{B}$ which is the modular matrix inverse (MMI) of $\boldsymbol{A}$*in* $Z_p$. Then,

$$\mathrm{mod}(\boldsymbol{B} \times \boldsymbol{A}, p) = \mathrm{mod}(\boldsymbol{A} \times \boldsymbol{B}, p) = \boldsymbol{I}_{(N+2)\times(N+2)} \tag{3}$$

where $\boldsymbol{I}_{(N+2)\times(N+2)}$ is the identity matrix of size $(N+2)\times(N+2)$. Matrix MMI($\boldsymbol{A}$) is calculated using Gauss Jordan Matrix Inverse in$Z_p$ [10]. Matrix $\boldsymbol{B}$ of size $(N + 2) \times (N + 2)$ is rewritten as,

$$\boldsymbol{B} = \mathrm{MMI}(\boldsymbol{A}, p) = \boldsymbol{A}^{-1} \mathrm{in} Z_p \tag{4}$$

Here, $\boldsymbol{B} \in Z_p^{(N+2)\times(N+2)}$. When there is no ambiguity, MMI($\boldsymbol{A}, p$) is simply written as $\boldsymbol{A}^{-1}$. Thus, $\boldsymbol{B}$ is calculated by the PKG as,

$$\boldsymbol{B} = \boldsymbol{A}^{-1} \tag{5}$$

Then in$Z_p$,

$$\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{I}_{(N+2)\times(N+2)} \tag{6}$$

The elements of $\boldsymbol{A}$ should be chosen such that $\boldsymbol{A}^{-1}$ exists. Now, USK matrix is $\boldsymbol{B} = A^{-1}$.

**Example 1** Let $p = 101$ and $N = 3$. Matrix $\boldsymbol{A}$ is randomly chosen as shown below.

$$\boldsymbol{A} = \begin{bmatrix} 44 & 34 & 63 & 79 & 51 \\ 3 & 21 & 53 & 86 & 7 \\ 55 & 62 & 14 & 50 & 43 \\ 44 & 30 & 52 & 85 & 10 \\ 43 & 27 & 19 & 8 & 13 \end{bmatrix}$$

Then, the calculated value of $\boldsymbol{B} = \boldsymbol{A}^{-1}$ is given below.

$$\boldsymbol{B} = \boldsymbol{A}^{-1} = \begin{bmatrix} 30 & 42 & 31 & 4 & 26 \\ 76 & 89 & 99 & 30 & 57 \\ 11 & 90 & 0 & 69 & 34 \\ 4 & 75 & 39 & 73 & 54 \\ 74 & 89 & 33 & 4 & 8 \end{bmatrix}$$

Here, it can be verified that mod($\mathbf{A} \times \mathbf{B}$, $p$) $= \boldsymbol{I}_{5\times 5.}$

## 3.6 *Relation Between the Rows of* A *and Columns of* B

Let the column vectors of B be written as,

$$\boldsymbol{B} = \left[\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_N, \boldsymbol{b}_{N+1}, \mathbf{b}_{N+2}\right] \tag{7}$$

Let the rows of matrix $\boldsymbol{A}$ be denoted by $\boldsymbol{a}_1$, $\boldsymbol{a}_2$,…,$\boldsymbol{a}_{\mathrm{i}}$,…, $\boldsymbol{a}_{N+2}$ as,

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ . \\ \boldsymbol{a}_i \\ . \\ \boldsymbol{a}_N \\ \boldsymbol{a}_{N+1} \\ \boldsymbol{a}_{N+2} \end{bmatrix} \tag{8}$$

Then (6) can be expressed as,

$$\boldsymbol{A} \times \boldsymbol{B} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ . \\ \boldsymbol{a}_i \\ . \\ \boldsymbol{a}_N \\ \boldsymbol{a}_{N+1} \\ \boldsymbol{a}_{N+2} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{b}_1 \; \boldsymbol{b}_2 \; . \; \boldsymbol{b}_j \; . \; \boldsymbol{b}_N \; \boldsymbol{b}_{N+1} \; \boldsymbol{b}_{N+2} \end{bmatrix} = \boldsymbol{I}_{(N+2)\times(N+2)} \tag{9}$$

From (9), it can be seen that, for $i = 1$ to $N + 2$ and $j = 1$ to $N + 2$, the product $\boldsymbol{a}_i \times \boldsymbol{b}_j$ is,

$$\boldsymbol{a}_i \times \boldsymbol{b}_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{10}$$

Thus, matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are orthogonal to each other in $Z_p$.

## *3.7 User Secret Keys*

User's secret keys (USKs) are the first $N$ column vectors of matrix $\boldsymbol{B}$, represented by $\boldsymbol{B}(1{:}N)$ as,

$$\boldsymbol{B}(1 : \boldsymbol{N}) = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_N] \tag{11}$$

Then the secret key of user$u_j$ is $\boldsymbol{b}_j$ for$j = 1$ to $N$. The size of $\boldsymbol{b}_j$ is $(N + 2) \times 1$ and $\boldsymbol{b}_j \in \mathrm{Z}_p^{(N+2)x1}$. Thus, the USK of the $j^{\text{th}}$ user$u_j$ denoted by $\text{USK}_j$ is $\boldsymbol{b}_j$. That is,

$$\text{USK}_j = \boldsymbol{b}_j \tag{12}$$

The USK values are also stored in UIR table for reference as shown in Table 1. The individual USKs are conveyed to the corresponding users through secured channels as shown in Fig. 1.

## *3.8 User Signature Verification Keys*

The signature verification key of user $u_j$ denoted by $\mathrm{VK}_j$ and is given by,

$$\mathrm{VK}_j = S \tag{13}$$

For $j = 1$ to $N$. Here, $\mathrm{VK}_j$ is same for all $j$. Verification keys $\mathrm{VK}_j$s are also conveyed to the corresponding users through secured channels as shown in Fig. 1.

## *3.9 Revocation List*

Revocation list (EL) is a binary row vector of size $1 \times N$ where $N$ is the total number of users. EL is represented by the row vector $\boldsymbol{E}$ as,

$$E = [e_1, e_2, \ldots, e_N] \tag{14}$$

The $j$th element $e_j$ stipulates whether user $u_j$ is revoked or not as follows.

$$e_j = \begin{cases} 0 \text{ means, } u_j \text{is revoked} \\ 1 \text{ means, } u_j \text{is enabled} \end{cases} \tag{15}$$

For $j = 1$ to $N$. Thus, vector $\boldsymbol{E}$ can be treated as the **Enable List** which represents the revoked users as given by (15). The size of $\boldsymbol{E}$ is $1 \times N$. In our scheme, the plaintext is encrypted such that the decrypted value for a revoked user is zero, whereas an unrevoked (enabled) user recovers the correct plain text after decryption. Thus, when $u_j$ is revoked, the decrypted result of $u_j$ is zero whatever may be the original plaintext value. Therefore, a revoked user fails to decrypt correctly. On the other hand, an unrevoked user can correctly decrypt the ciphertext. For proper management, EL is also stored in UIR as shown in Table 1.

### 3.9.1 Broadcast Data to be Encrypted

Here, the plaintext to be encrypted and broadcasted is an integer vector $\boldsymbol{Q}$ of size $N \times 1$, (it can be a set of secret keys for certain symmetric encrypted files). The elements of $\boldsymbol{Q}$ should be in the range $1 \le q \le (p - 1)$. That is, $\boldsymbol{Q} \in \boldsymbol{Z}_{\boldsymbol{p}}^{N \times 1}$. All zero message vector $\boldsymbol{Q}$ is avoided as it is reserved to represent the result of decryption for revoked users.

### 3.9.2 Encryption Preprocess Matrix

The encryption preprocess matrix (EPM), represented by $\boldsymbol{F}$, is formulated by extending the weighted Revocation list $\boldsymbol{E}$ to include the signature vector $\boldsymbol{S}$ and a randomization vector $\boldsymbol{R}$ as,

$$F = [Q \times E,\ S,\ R] \tag{16}$$

Here, $\boldsymbol{S}$ is the signature vector and $\boldsymbol{R}$ is the randomizing vector. Matrix $\boldsymbol{F}$ is obtained by concatenating vector $\boldsymbol{S}$ and vector $\boldsymbol{R}$ to$\mathbf{Q}$ $\times\mathbf{E}$.

### 3.9.3 Signature Vector

The signature vector $\boldsymbol{S}$ belongs to$Z_p$. The size of $\boldsymbol{S}$ is $N \times 1$, and it is sent to all the users from PKG, at the start of the session as shown in Fig. 1. The value of $\boldsymbol{S}$ remains same in successive encryptions.

### 3.9.4 Randomization Vector

The randomization vector $\boldsymbol{R}$ of size $N \times 1$ belongs to $Z_p$. It varies randomly from encryption to encryption. This randomization of ciphertext prevents chosen ciphertext attack as will be explained later.

### 3.9.5 Matrix F

The size of matrix $\boldsymbol{F}$ is $N \times (N + 2)$. In (16), the size of $Q \times \boldsymbol{E}$ is $(N \times 1) \times (1 \times N) = N \times N$. The last two columns of $F$ are $S$ and $R$, respectively. To make it clear, let matrix $\boldsymbol{F}$ be represented in terms of its column vectors as,

$$\boldsymbol{F} = \left[\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N, \boldsymbol{f}_{N+1}, \boldsymbol{f}_{N+2}\right] \tag{17}$$

Then,

$$\left[\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N\right] = Q \times \boldsymbol{E} = [Q \times e_1, Q \times e_2, \ldots, Q \times e_{\mathrm{N}}] \tag{18}$$

$$\boldsymbol{f}_{N+1} = \boldsymbol{S} \tag{19}$$

$$\boldsymbol{f}_{N+2} = \boldsymbol{R} \tag{20}$$

Here, the size of $\boldsymbol{f}_i$ is $N \times 1$, for $i = 1$ to $N + 2$.

# 4 Identity-Based Broadcast Encryption

The basic operations of IBBE-MK are described in this section.

## *4.1 Encryption*

The encryptor (owner of data) generates the ciphertext $\boldsymbol{C}$ which is a matrix of size $N \times (N + 2)$as,

$$\boldsymbol{C} = \ \mathrm{mod}(\boldsymbol{F} \times \boldsymbol{A}, p) \ = \boldsymbol{F} \times \boldsymbol{A} \tag{21}$$

In terms of the column vectors of F and row vectors of $\boldsymbol{A}$, Eq. (21) can be expanded as,

$$\boldsymbol{C} = \boldsymbol{f}_1 \times \boldsymbol{a}_1 + \ldots + \boldsymbol{f}_N \times \boldsymbol{a}_N + \boldsymbol{f}_{N+1} \times \boldsymbol{a}_{N+1} + \boldsymbol{f}_{N+2} \times \boldsymbol{a}_{N+2} \tag{22}$$

## *4.2 Extraction Property of* C

Equation (22) can be rewritten as,

$$\boldsymbol{C} = \sum_{i=1}^{N+2} \boldsymbol{f}_i \times \boldsymbol{a}_i \tag{23}$$

Consider the product $\boldsymbol{C} \times \boldsymbol{b}_j$ for $j = 1$ to $N + 2$. From (23),

$$\boldsymbol{C} \times \boldsymbol{b}_j = \sum_{i=1}^{N+2} \boldsymbol{f}_i \times \boldsymbol{a}_i \times \boldsymbol{b}_j \tag{24}$$

Equation (24), over the range $i = 1$ to $(N + 2)$ can be split as,

$$\boldsymbol{C} \times \boldsymbol{b}_j = \sum_{i=1, i \neq j}^{N+2} \boldsymbol{f}_i \times \boldsymbol{a}_i \times \boldsymbol{b}_j + \sum_{i=j} \boldsymbol{f}_i \times \boldsymbol{a}_i \times \boldsymbol{b}_j \tag{25}$$

When $i \neq j$, from (10), $\boldsymbol{a}_i \times \boldsymbol{b}_j = 0$. Therefore, the first summation on the RHS of (25) is zero. When $i = j$, from (10), $\boldsymbol{a}_i \times \boldsymbol{b}_j = 1$. Therefore, the second term on the RHS of (25) is $\boldsymbol{f}_j$. Hence, Eq. (25) leads to

$$\boldsymbol{C} \times \boldsymbol{b}_j = \boldsymbol{f}_j \tag{26}$$

For $j=$ 1 to $(N + 2)$. Thus, $f_j$ can be extracted from $\boldsymbol{C}$ using (26). Ciphertext $\boldsymbol{C}$ is sent to user $u_j$ on his/her request.

## *4.3 Signature Verification*

On receiving $\boldsymbol{C}$, user $u_j$ calculates the received signature as,

$$\boldsymbol{Srec} = \bmod(\boldsymbol{C} \times \boldsymbol{b}_{N+1}, p) = \boldsymbol{C} \times \boldsymbol{b}_{N+1} \tag{27}$$

Assuming no error (or alterations) in received $\boldsymbol{C}$, from (26) and (27), we get,

$$\boldsymbol{Srec} = \boldsymbol{f}_{N+1} \tag{28}$$

From (19) and (28),

$$\boldsymbol{Srec} = \boldsymbol{S} \tag{29}$$

Thus, when ciphertext $\boldsymbol{C}$ is not tampered, condition $\boldsymbol{Srec} = \boldsymbol{S}$ is satisfied and the signature verification is successful.

Condition $\boldsymbol{Srec} \neq \boldsymbol{S}$ means the signature verification has failed. This implies error or tampering in the received ciphertext $\boldsymbol{C}$. Then the receiver discards $\boldsymbol{C}$ and sends a repeat request. The decryptor at the receiver is designed such that it indicates whether the signature verification is successful or not.

## *4.4 Decryption*

After successful signature verification, user $u_j$ recovers the message vector, using his secret key $\boldsymbol{b}_j$ as,

$$\boldsymbol{d}_j = \boldsymbol{C} \times \boldsymbol{b}_j \tag{30}$$

for $j = 1$ to $N$. From (30) and (26),

$$\boldsymbol{d}_j = \boldsymbol{f}_j \tag{31}$$

Here, the size of vector $\boldsymbol{d}_j$ is $N \times 1$.
From (18) and (31),

$$\boldsymbol{d}_j = \boldsymbol{Q} \times e_j \tag{32}$$

When user $u_j$ is revoked, $e_j = 0$ and then from (32), $\boldsymbol{d}_j = 0$, else.
$e_j = 1$ and $\boldsymbol{d}_j = \boldsymbol{Q}$. Thus, $\boldsymbol{d}_j$ given by (30) can be expressed as,

$$\boldsymbol{d}_j = \boldsymbol{C} \times \boldsymbol{b}_j = \begin{cases} 0 \text{ if } u_j \text{ is revoked} \\ \boldsymbol{Q} \text{ if } u_j \text{ is enabled} \end{cases} \tag{33}$$

Equation (33) holds good for $j = 1$ to $N$. In this way, the unrevoked (enabled) user recovers original plaintext vector $\boldsymbol{Q}$.

**Example 2** Matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and modulus $p$ are same as in Example 1. Let plaintext vector $\boldsymbol{Q}$ be $[11, 68, 9]^{\mathrm{T}}$ and the signature vector $\boldsymbol{S}$ be $[60, 23, 11]^{\mathrm{T}}$. Here $N = 3$.

Case1: Revocation list is taken as $\boldsymbol{E} = [1]$. That is, all the users are enabled. Taking the randomizing vector $\boldsymbol{R}$ as
$\boldsymbol{R} = [75, 98, 36]^{\mathrm{T}}$, the Encryption Preprocess Matrix $\boldsymbol{F}$ is obtained using (16) as,

$$\boldsymbol{F} = \begin{bmatrix} 11 & 11 & 11 & 60 & 75 \\ 68 & 68 & 68 & 23 & 98 \\ 19 & 19 & 19 & 11 & 76 \end{bmatrix}$$

Ciphertext $\boldsymbol{C}$ is calculated in $Z_p$ and found to be,

$$\boldsymbol{C} = \boldsymbol{F} \times \boldsymbol{A} \begin{bmatrix} 18 & 62 & 16 & 86 & 60 \\ 42 & 81 & 81 & 88 & 90 \\ 34 & 60 & 42 & 73 & 88 \end{bmatrix}$$

Let there be no tampering in $\boldsymbol{C}$. Then the received signature is calculated (in modular arithmetic) using (27) as,

$$\boldsymbol{Srec} = \boldsymbol{C} \times \boldsymbol{b}_4 = \begin{bmatrix} 18 & 62 & 16 & 86 & 60 \\ 42 & 81 & 81 & 88 & 90 \\ 34 & 60 & 42 & 73 & 88 \end{bmatrix} \times \begin{bmatrix} 4 \\ 30 \\ 69 \\ 73 \\ 4 \end{bmatrix} = \begin{bmatrix} 60 \\ 23 \\ 11 \end{bmatrix}$$

Since $\boldsymbol{Srec} = \boldsymbol{S}$, the signature verification is successful.
Decryption at $u_1$ using (30) yields,

$$\boldsymbol{d}_1 = \boldsymbol{C} \times \boldsymbol{b}_1 = \begin{bmatrix} 18 & 62 & 16 & 86 & 60 \\ 42 & 81 & 81 & 88 & 90 \\ 34 & 60 & 42 & 73 & 88 \end{bmatrix} \times \begin{bmatrix} 30 \\ 76 \\ 11 \\ 4 \\ 74 \end{bmatrix} = \begin{bmatrix} 11 \\ 68 \\ 9 \end{bmatrix}$$

Similarly, it can be shown that $\boldsymbol{d_2} = \boldsymbol{d_3} = \boldsymbol{d_1} = \boldsymbol{Q}$.
Case 2: Revocation list is taken as, $\boldsymbol{E} = [1, 0, 1]$. User $u_2$ is revoked. In this case,

$$\boldsymbol{F} = \begin{bmatrix} 11 & 0 & 11 & 60 & 75 \\ 68 & 0 & 68 & 23 & 98 \\ 19 & 0 & 19 & 11 & 76 \end{bmatrix}$$

$$\boldsymbol{C} = \boldsymbol{F} \times \boldsymbol{A} = \begin{bmatrix} 86 & 33 & 39 & 49 & 84 \\ 40 & 67 & 12 & 98 & 18 \\ 78 & 65 & 45 & 55 & 56 \end{bmatrix}$$

Here, signature verification is successful same as in case 1. Decryption at $u_1$, $u_2$, and $u_3$ using (33) in modular arithmetic yields,

$$\boldsymbol{d_1} = \boldsymbol{C} \times \boldsymbol{b}_1 = \begin{bmatrix} 11 \\ 68 \\ 9 \end{bmatrix}, \boldsymbol{d_2} = \boldsymbol{C} \times \boldsymbol{b}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \boldsymbol{d_3} = \boldsymbol{C} \times \boldsymbol{b}_3 = \begin{bmatrix} 11 \\ 68 \\ 9 \end{bmatrix}$$

Now, data $d_2$ is all zeros which implies that User $u_2$ is revoked.

## 5 Basic Layout of IBBE-MK

The basic layout of the proposed IBBE-MK is shown in Fig. 2. In Fig. 2, the data owner is also the data encryptor. The encrypted data is stored in the semi-trusted cloud storage. The users can access the ciphertext after proper authentication. In Fig. 2, the authenticator is a part of the proxy server.

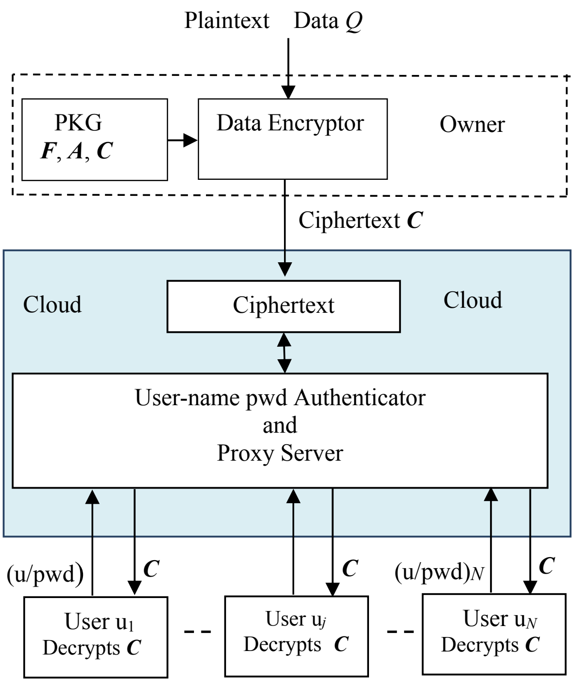



**Fig. 2** Basic layout of IBBE-MK

### 5.1 Authentication, Downloading and Decryption

User authentication is provided by user-name/password ($u$/pwd) method. User $u_j$ submits his ($u$/pwd)$_j$ to the authenticator as shown in Fig. 2. After verification by the authenticator, the proxy server allows User $u_j$ to download the ciphertext $\boldsymbol{C}$. User $u_j$ in turn verifies the signature and decrypts $\boldsymbol{C}$ and recovers $\boldsymbol{d}_j$ as explained earlier.

### 5.2 Ciphertext Expansion Ratio

Ciphertext expansion ratio (CER) is the ratio of the size of the ciphertext to that of its plaintext. A higher value of CER means the computational cost, and communication overheads are higher. Smaller the CER, better is the encryption efficiency. CER is defined as,

$$\mathrm{CER} = \frac{\text{Size of Ciphertext}}{\text{Size of Plaintext}} \tag{34}$$

In our IBBE-MK, the size of the ciphertext $\boldsymbol{C}$ is $N \times (N + 2)$, and the size of the plaintext $\boldsymbol{Q}$ is $N \times 1$.

Therefore,

$$\mathrm{CER} = \frac{N \times (N+2)}{N} = N + 2 \tag{35}$$

## 5.3 Security Analysis of IBBE-MK

Brute Force Attack: It is almost infeasible for an attacker to guess the secret key of the user because the probability of correct guessing is $p^{-(N+2)}$ where the size of the key is $(N + 2)$ and $p$ is the number of ways in which an integer can be selected in$Z_p$.

Collusion Resistance: Even when all the users collude, the resulting secret key matrix has columns $\boldsymbol{b}_1$ to $\boldsymbol{b}_{N+1}$ but not $\boldsymbol{b}_{N+2}$. Hence, the resulting matrix is of size $(N + 2) \times (N + 1)$ and not a square matrix, and it has no inverse to recover the MSK matrix $\boldsymbol{A}$. Also, several known user secret keys put together cannot recover another secret key as they are linearly independent.

Chosen Plaintext Attack: In chosen plaintext attack (CPA), the attacker chooses his plaintexts according to his choice, gets the corresponding ciphertexts and based on these observations, he seeks to calculate the secret key(s) of the encryptor. In IBBE-MK, the encryptions are randomized using different random integer values for $\boldsymbol{R}$ for successive encryptions. Thus, the value of $\boldsymbol{R}$ is chosen fresh for each consecutive plaintext. Thus, for the same plaintext, the ciphertext would be different in successive trials. Therefore, IBBE-MK provides protection against CPA.

Chosen Ciphertext Attack: Chosen ciphertext attack (CCA) attempts to break the security of the decryption process. The attacker chooses different ciphertexts, gets the corresponding decrypted plaintexts and using them the attacker strives to crack the decryption key(s). In IBBE-MK, the digital signature $\boldsymbol{S}$ embedded in Encryption Preprocess Matrix $\boldsymbol{F}$ provides protection against CCA. In IBBE-MK, the decryptor evaluates each ciphertext input for valid signature $\boldsymbol{S}$. If the signature verification fails, the decryptor outputs zero as the plaintext. Therefore, the CCA attacker cannot crack the decryptor as most of the plaintext outputs are empty (zeros).

## 5.4 Comparison of IBBE-MK with Other Methods

In this section, performance characteristics of IBBE-MK are compared with the existing methods by Agrawal, Boneh, Boyen (ABB) [20], Gentry, Halevi, Vaikuntanathan (GHV) [21], Wang, Liu Wang (WLW) [23] and Wang, Bi (WB) [24]. In a crypto-system, a low CER implies efficient encryption. Also, shorter the key better is the performance. The comparative values of the lengths of plaintext, ciphertext and secret key as well as CER are shown in Table 2.

Here, $n$ = security parameter that represents the dimension of the lattice. In IBBE-MK, $n = N$ itself. $p$ = modulus of the finite field $Z_p$ and $u = \mathrm{ceil}(\log_2(p))$ = Upper bound on the size of an integer in $Z_p$. Parameter $k$ = Actual number of receivers

**Table 2** Comparison of lengths of various parameters and CER

| Methods | Message (Plaintext) length in bits | Ciphertext length in bits | Length of USK in bits per bit of message | Ciphertext expansion ratio (CER) |
|---|---|---|---|---|
| IBBE-MK | $n*u$ | $n*(n+2)*u$ | $(n+2)/n$ | $(n+2)$ |
| ABB [20] | 1 | $(2*(2*n*u)+1)*u$ | $n*u$ | $(2*(2*n*u)+1)*u$ |
| GHV [21] | $(2*n*u)*(2*n*u)$ | $(2*n*u)*(2*n*u)*u$ | $u$ | $u$ |
| WLW [23] | 1 | $((2*n*u)+1)*u$ | $(2*n*u)*u$ | $((2*n*u)+1)*u$ |
| WB [24] | $u$ | $(k+2)*u$ | $(k+1)*(2*n*u)$ | $(k+2)$ |

**Table 3** Comparison of computational cost

| Methods | Computational cost in terms of no. of bit multiplications per bit of message | |
|---|---|---|
| | Encryption | Decryption |
| IBBE-MK | $n*u+(n+2)^2*u$ | $((n+2)^2/n)*u$ |
| ABB [20] | $(2*(2*n*u)+1)*n*u^2$ | $2*(2*n*u)*u^2$ |
| GHV [21] | $n*u^2$ | $2*(2*n*u)*u^2$ |
| WLW [23] | $((2*n*u)+1)*n*u^2$ | $(2*n*u)*u^2$ |
| WB [24] | $(k+1)*(2*n*u)*n*u$ | $(k+1)*(2*n*u)*u$ |

(users). $k$ should be less than or equal to $n$ for correct broadcast encryption. In the present context, $k$ is set equal to $N$. Each integer multiplication requires $u^2$ bit multiplications.

From Table 2, it can be seen that CER of IBBE is reasonably low compared to ABB [20] and WLW [23].

Computational costs of encryption and decryption for each method are shown in Table 3. Computational cost is expressed in terms of number of bit multiplications per bit of the plaintext (message).

Another feature of IBBE-MK is the computational cost of decryption is the lowest compared to other four methods. The entire operation of IBBE-MK is substantially simple compared to the other lattice-based methods.

## 6 Conclusion

A new method IBBE-MK for Identity-Based Broadcast Encryption using matrix keys is presented. The novelty of IBBE-MK is, it provides integrated signature verification and randomization with short ciphertext. The revocation is direct and controlled by the owner of data. The revocation and inclusion are carried out efficiently by just changing the bits of the revocation vector. Apart from easy revocation/inclusion, IBBE-MK can accommodate new users by initially using large sized master secret key and associated secret keys. The security level of IBBE-MK can be increased by increasing the modulus value $p$ and the length of the secret keys at the cost of moderate increase in CER and computational costs of encryption and decryption.

## References


1. Shamir A (1985) Identity-based cryptosystems and signature schemes. In: Blakley GR, Chaum D (eds) Advances in cryptology CRYPTO 1984 LNCS, vol 196. Springer, Berlin, Heidelberg, pp 47–53. https://doi.org/10.1007/3-540-39568-7_5
2. Fiat A, Naor M (1994) Broadcast encryption. In: Stinson DR (eds) Advances in cryptology CRYPTO 1993 LNCS, vol 773. Springer, Berlin, Heidelberg, pp 480–491. https://doi.org/10.1007/3-540-48329-2_40
3. Delerablée C (2007) Identity-based broadcast encryption with constant size ciphertexts and private keys. In: Kurosawa K (ed) Advances in cryptology ASIACRYPT 2007 LNCS, vol 4833. Springer, Berlin, Heidelber, pp 200–215. https://doi.org/10.1007/978-3-540-76900-2_12
4. Attrapadung N, Imai H (2009) Attribute-based encryption supporting direct/indirect revocation modes. In: Parker MG (ed) Cryptography and coding IMACC 2009 LNCS, vol 5921. Springer, Berlin, Heidelberg, pp 278–300. https://doi.org/10.1007/978-3-642-10868-6_17
5. Boldyreva A, Goyal V, Kumar V (2008) Identity-based encryption with efficient revocation. In: Proceedings of the 15th ACM conference on computer and communications security CCS 2008 ACM. Alexandria, Virginia, USA, pp 417–426. https://doi.org/10.1145/1455770.1455823
6. Boneh D, Ding X, Tsudik G, Wong M (2001) A method for fast revocation of public key certificates and security capabilities. In: Proceedings of the 10th conference on USENIX security symposium SSYM 2001 ACM, vol 10. Washington D.C, USA, p 22. https://doi.org/10.5555/1251327.1251349
7. Libert B, Quisquater J-J (2003) Efficient revocation and threshold pairing based cryptosystems. In: Proceedings of the twenty-second annual symposium on principles of distributed computing PODC 2003 ACM. Boston, Massachusetts, USA, pp 163–171. https://doi.org/10.1145/872035.872059
8. Attrapadung N, Libert B (2010) Functional encryption for inner product: achieving constant-size ciphertexts with adaptive security or support for negation. In: Nguyen PQ, Pointcheval D (eds) Public Key cryptography PKC 2010 LNCS, vol 6056. Springer, Berlin, Heidelberg, pp 384–402. https://doi.org/10.1007/978-3-642-13013-7_23
9. Ramanna SC (2016) More efficient constructions for inner-product encryption. In: Manulis M, Sadeghi AR, Schneider S (eds) Applied cryptography and network security ACNS 2016 LNCS, vol 9696. Springer, Cham, pp 231–248. https://doi.org/10.1007/978-3-319-39555-5_13
10. Sakai R, Furukawa J (2007) Identity-based broadcast encryption. Cryptology ePrint Archive, Report 2007/217. https://eprint.iacr.org/2007/217.pdf. Accessed on July 28, 2020
11. Hur J, Park C, Hwang SO (2012) Privacy-preserving identity-based broadcast encryption. Inform. Fusion 13(4):296–303. https://doi.org/10.1016/j.inffus.2011.03.003
12. Lai J, Mu Y, Guo F, Susilo W, Chen R (2016) Anonymous identity-based broadcast encryption with revocation for file sharing. In: Liu J, Steinfeld R (eds) Information security and privacy ACISP 2016 LNCS, vol 9723. Springer, Cham, pp 223–239. https://doi.org/10.1007/978-3-319-40367-0_14
13. Boneh D, Gentry C, Waters B (2005) Collusion resistant broadcast encryption with short ciphertexts and private keys. In: Shoup V (eds) Advances in cryptology—CRYPTO 2005 LNCS, vol 3621. Springer, Berlin, Heidelberg, pp. 258–275. https://doi.org/10.1007/11535218_16
14. Zhu Y, Yu R, Chen E, Huang D (2018) Dual-mode broadcast encryption. Sci China Inf Sci 61:118101. https://doi.org/10.1007/s11432-017-9287-6

15. Ge A, Wei P (2019) Identity-based broadcast encryption with efficient revocation. In: Lin D, Sako K (eds) Public-Key cryptography—PKC 2019 LNCS, vol 11442. Springer, Cham, pp 405–435. https://doi.org/10.1007/978-3-030-17253-4_14
16. Boneh D, Hamburg M (2008) Generalized identity based and broadcast encryption schemes. In: Pieprzyk J (ed) Advances in cryptology—ASIACRYPT 2008 LNCS, vol 5350. Springer, Berlin, Heidelberg, pp 455–470. https://doi.org/10.1007/978-3-540-89255-7_28
17. Wang XA, Yang X (2009) Identity based broadcast encryption based on one to many identity based proxy re-encryption. In: 2nd IEEE international conference on computer science and information technology, Beijing, 47–50 2009. https://doi.org/10.1109/ICCSIT.2009.5234998
18. Li D, Liu J, Zhang Z, Wu Q, Liu W (2018) Revocable hierarchical identity-based broadcast encryption. Tsinghua Sci Tech 23(5):539–549. https://doi.org/10.26599/TST.2018.9010023
19. Susilo W, Chen R, Guo F, Yang G, Mu Y, Chow Y-W (2012) Recipient revocable identity-based broadcast encryption: how to revoke some recipients in IBBE without knowledge of the plaintext. In: Proceedings of the 11th ACM on Asia conference on computer and communications security ACM. Xi'an, China, pp 201–210. https://doi.org/10.1145/2897845.2897848
20. Agrawal S, Boneh D, Boyen X (2010) Efficient lattice (H)IBE in the standard model. In: Gilbert H (ed) Advances in cryptology—EUROCRYPT 2010 LNCS, vol. 6110. Springer, Berlin, Heidelberg, pp 553–572. https://doi.org/10.1007/978-3-642-13190-5_28
21. Gentry C, Halevi S, Vaikuntanathan V (2010) A simple BGN-Type cryptosystem from LWE. In: Gilbert H (ed) Advances in cryptology—EUROCRYPT 2010 LNCS, vol 6110. Springer, Berlin, Heidelberg, pp 506–522. https://doi.org/10.1007/978-3-642-13190-5_26
22. Gentry C, Peikert C, Vaikuntanathan V (2008) Trapdoors for hard lattices and new cryptographic constructions. In: Proceedings of the fortieth annual ACM symposium on Theory of computing, ACM. Victoria, British Columbia, Canada, pp 197–206. https://doi.org/10.1145/1374376.1374407
23. Wang F, Liu Z, Wang C (2016) Full secure identity-based encryption scheme with short public key size over lattices in the standard model. Int J Comput Math 93(6):854–863
24. Wang J, Bi J (2018) Lattice-based identity-based broadcast encryption. In: IACR Eprint archive, ePrint 2010/288.
25. Zhang X, Wang S, Zhang W (2015) Forward-secure identity-based broadcast encryption scheme from lattice. Appl Math Inf Sci 9(4):1993–2000

# Sentiment Analysis to Detect Depression in Social Media Users: Overview and Proposed Methodology

**P. Ushashree, G. Harshika, Umme Haani, and Rishabh Kalai**

## 1 Introduction

Despite the fact that it is quite a prevalent mental health disorder, depression is the most under-diagnosed disorder due to the lack of proper testing and the social stigma surrounding it. The WHO gives an estimate that nearly 264 million people under all age groups suffer from depression. The WHO ranks depression as one of the most burdensome diseases in the world [1].

Depression has a wide range of symptoms which include symptoms of anxiety, disturbed sleep. It also includes hyper insomnia and insomnia and loss of appetite, guilt, low self-worth, and loss of interest in activities that were enjoyable. The symptoms worsen depending on the severity of the disorder and may lead to bipolar affective disorder (BAD). Depression is also known to affect other present chronic conditions like cancer, diabetes, and cardiovascular diseases along with affecting the social life of an individual and interpersonal relationships [2]. Severe depression may also lead to suicide [3].

Diagnosis and proper treatment of depression can greatly improve the condition; however, it is quite di cult to diagnose it at a population-level, and a tedious, time-consuming, and expensive task. Under such circumstances, social media provides a treasure trove of behavioral information holding data relevant to the detection of depression in users. Social media depicts attributes that are relevant in capturing the mood and thinking patterns, communication, and socialization. People suffering from depression tend to withdraw themselves from social situations and activities. Their emotions and language can be traced via their posts which may indicate feelings of

P. Ushashree (✉) · G. Harshika · U. Haani
Department of CSE, Nitte Meenakshi Institute of Technology, Bangalore, India
e-mail: ushashree.sgs@nmit.ac.in

R. Kalai
Department of CSE, BNM Institute of Technology, Bangalore, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_39

hopelessness, guilt, suicidal expressions, self-hatred, and helplessness. These are all indicators of a severe onset of depression [2].

Twitter is one such social media platform that allows its users to micro-blog with up to 120 characters per blog. It has a wide variety of topics ranging from politics to health. It is deemed to be most appropriate for research in the mental health domain as mental health research requires study related to behavior which includes the way they interact and engage with other users online including their friends and family. Research via the capture of population-level behavioral trends from online data has been used by health researchers successfully [4]. Several studies on mental health and depression detection have already been performed using Twitter data as the predominant data source [4]. The extraction of such data is called sentiment analysis, which is basically a tool to gauge the attitude, emotion, sentiment, and opinions of an individual through information written by them. More details on sentiment analysis and its use cases can be found in [5]. However, even with extensive work done on this subject, there are several loopholes that make this method of detection unstable if it were used on a larger scale. In our review we aim to answer the following questions:

*RQ1*: Is there any established framework or methodology to detect depression through Twitter analysis?
*RQ2*: Is there any way to improve the quality of data and make the data more reliable?

In order to investigate these questions, we have conducted a Systematic Literature Review (SLR) in hopes of improving the currently available technology to make it more user-friendly, reliable, and provide greater accuracy.

## 2 Related Work

### *2.1 Background Literature*

Throughout history, several online social networks of individuals have been studied in order to analyze behavioral changes related to the psychological environments. The roles of stress in social situations and methods to cope with them for individuals undergoing treatment for depression have been studied by Billings and Moo. An intake of less than 500 kcal/day can lead to glycogen stores be-ing mobilized in the body and 24 h fasting can cause lipolysis and accelerated protein catabolism [6]. Studies have suggested that fasting in order to decrease body weight can facilitate depression. Restricting calories may impair cognitive abilities and lower overall quality of life resulting in stress and an elevation of the level of corticosterone [7, 8] which may lead to an increased risk for depression.

The correlation between sleep and behavior exhibited by a person has been extensively researched over the years. For instance, people suffering from insomnia are prone to having higher levels of depression and anxiety than those with normal sleep schedules. They are 10 times as likely to have clinical depression and 17 times as likely to have clinical anxiety than the average person with a regular sleep schedule.

Oxman [9] showed through his research that the linguistic analysis of speech could classify individuals into groups of those who were not suffering from depression and those who were. Similar results are obtained under a computerized analysis of a written text. This opens the door to utilizing social media for analyzing mental health. NLP can be applied to obtain relevant results and make predictions with around 80 percent accuracy [2, 10].

Twitter has been widely used in studies to show that individuals who score high in depression scales tend to have differences in writing which is quantifiable with LIWC. Individuals suffering from depression often make extensive use of depression-related words and emotions when compared to their normal counterparts [11].

## *2.2 Data Collection Techniques*

Upon survey of quite a few papers that use Twitter as a medium for detecting depression, the following trends have been observed with regards to the various methods used to collect relevant data. Some of the most prevalent ways are:

1. Crowd-sourcing: Crowd-sourcing [2, 10] is used in order to study the control group of users with verified history with depression or at its onset. This also helps to decrease the ethical concerns with studying such data and reliability issues with information collected, however, its issues with scalability leave more to be desired. This often leads to the application of supervised machine learning algorithms where all the required data is labeled and given in the form of a training data set to the system.
2. Twitter API: Twitter API [3, 12] is used in order to extract all public data ranging from a specific time period. This is further preprocessed by identifying those with emotion-related hashtags in order to extract relevant information. This makes most of the work extremely tedious and in order to create a control group, in most cases, human coders were required to verify the tweets manually. This leads to the implementation of semi-supervised machine learning algorithms where a small amount is labeled and combined with a large amount of unlabeled data and given as a training data set to the system.

## *2.3 Classification and Identification Techniques*

Among various techniques used to classify a user as depressed, the most relevant and predominant are Classification frameworks based on LIWC (Linguistic Inquiry and Word Count):

1. LIWC is a tool used to quantify data regarding the mental state of a patient from their writing. It has been extensively used in research regarding sentiment

analysis. This is further used to predict tendencies in psychiatric disorders and neurotic tendencies. [13] The data extracted from the tweets are labeled in accordance with LIWC standards mainly into words having a positive affect and those having a negative affect. These are further subdivided based on the experiment, usually to ag the level of negativity expressed by the rates explained in LIWC. [13, 14] These are then scored and individuals with a higher score with words categorized as negative, are tagged as either strongly concerning, possibly concerning, or safe to ignore [3].

2. Machine learning classifiers:

    (a) *Support Vector Machine (SVM) classifier:* This classifier is used in supervised learning for classification and regression. Twitter posts are represented as vectors of features which include a variety of conditions to t into a feature which has been highly researched as common behaviors among people dis-playing depression, like duration of use, as people suffering from depression tend to be active during the night [2], hashtags, language expression, emotional context and level of engagement and tokenized for feature extraction. Using this, the accuracy of more than 70 percent was achieved in predicting depression, by [10] and [15]. Since it is a supervised learning algorithm, all the required data and features were labeled and then provided to the training data set.

    (b) *Multinomial Naive Bayes (MNB) and Liblinear classifiers:* This classification system was used in conjunction where the MNB classifier outperformed Liblinear inaccuracy by [4]. MNB works extremely well with data that can be turned into counts. It gives the probability of the occurrence of a word or feature which is then used to determine the severity of the situation depending on LIWC. Liblinear is a linear classifier for data with huge amounts of features. Since it is an implementation of SVM, it works very similarly to it, wherein it gives the feature probabilities. Both these algorithms work very similar to SVM, however, the main reason for selecting these algorithms is because of their ability to scan and classify a massive number of Tweets which is essential in scaling the operation of detecting depression at the population level.

3. Human Coding: In this method, coders were asked to analyze and classify tweets as concerning and not concerning. These were further subdivided, and the implementation was done on a small sample of 100 tweets. This human coded data for classification was later used by machine classifiers as training data which could automatically detect the category of concern (Fig. 1).

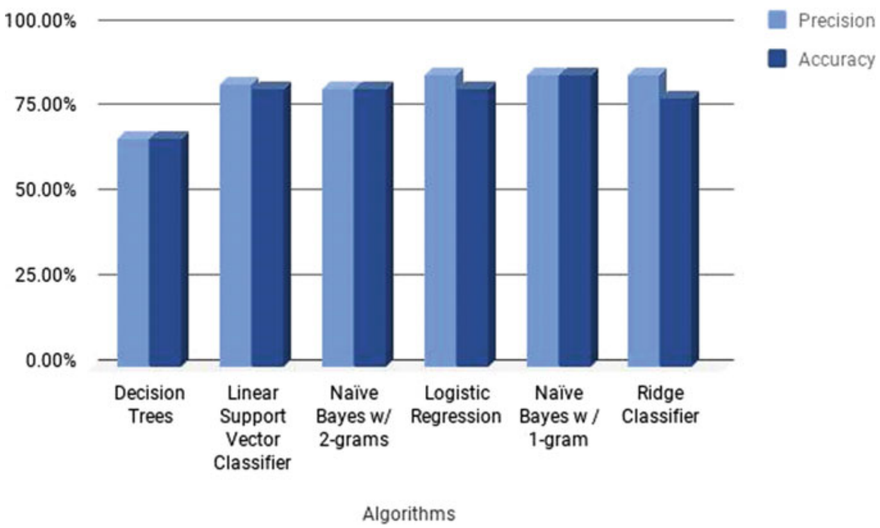



**Fig. 1** Precision and accuracy for different classification models [16]

## 2.4 Advantages

1. *Easier detection on population-level:* It provides a less expensive and less tedious larger scale detection system for depression that can be used to provide resources.
2. *Likelihood prediction*: Not only can it identify depression; it can also predict the likelihood of having depression in the future. This could highly improve the way mental health is treated and diagnosed in the future [10].
3. *Increased accuracy*: As the data-sets used for training increase, the accuracy of prediction also increases [4].
4. *Identification* via*ego-networks*: After detecting one individual, their ego network (a group of people they interact with the most) can also be analyzed and as shown by [10] This leads to faster rates of detection.

## 2.5 Disadvantages

1. *Ethical and privacy issues:* Ethical concerns like privacy issues come into the picture, as extracting information without the user's consent, even if it is public information. Most users suffering from depression would be reluctant to allow such extraction of information. This has been mentioned as a cause of concern in [2, 3, 10, 12]. Until there is a great deal of advancement in cyber-security, the privacy of any user cannot be ensured.

2. *Lack of sophisticated analysis:* Despite LIWC being a highly sophisticated language classification tool, the context of literature becomes irrelevant when the user is tweeting sarcastically. The context of the tweet matters a lot, and this leads to false positives; wasting valuable resources and time.
3. *Small test group size:* Most of the research done in the domain has used a rather small size of participants to perform the experiments. This makes the results obtained highly unreliable as usually drastic changes are observed on increasing the number of participants, features, or even the range of time during which the experiment is carried out. The accuracy provided by any of the models does not exceed 80 percent.
4. *Unreliable:* Reliability also comes under question when figuring if personality portrayed online actually relates to a person's real personality. Most people tend to showcase an ideal image of themselves on social media, and tweets can be deleted, which again leads to unrequired data as none of these methods focus on updating information that has already been collected.
5. *Lack of language identification:* The classifiers only work in English which ignores several languages that are also used on the platform to micro-blog. Some tweets are typed in English, but they are in other languages which are undoubtedly not recognizable by LIWC or any such tool.
6. *Lack of any standard framework for detection*: In order to make this approach readily accessible to medical professionals, there needs to be a standardized approach to determine the accuracy which is supported by other external factors of an individual. There is also a lack of a go-to reliable framework to increase ease of use. As of now, this method of detection is seemingly more tedious than traditional methods.

## 3 Methodology

To address some of the mentioned issues, we propose a methodology which follows a cumulative analysis approach. An approach that embodies various aspects of behavioral and characteristic changes a person portrays during depression, along with the results from the sentiment analysis.The two main observable behavioral changes a person shows are the changes in the sleep pattern [17, 18] and changes in their physical activities on a day to day basis.

During depression, people often suffer from various sleep abnormalities like insomnia and hyper insomnia [19, 20]. Data from research show that the majority of the population reported sleep difficulties during depression [21]. Sleep controls of depressed people are often impaired, the period of wakefulness is high, reducing sleep efficiency [22]. This can be identified using REM (rapid eye movements). During sleep, eye movements are higher during a disturbed sleep [23, 24] (Fig. 2).

Depression leads to loss of appetite and loss of interest to perform any activity; this is a data point that helps us gain an association between physical activities taking place in different contexts and the symptoms of depression. Research shows that the

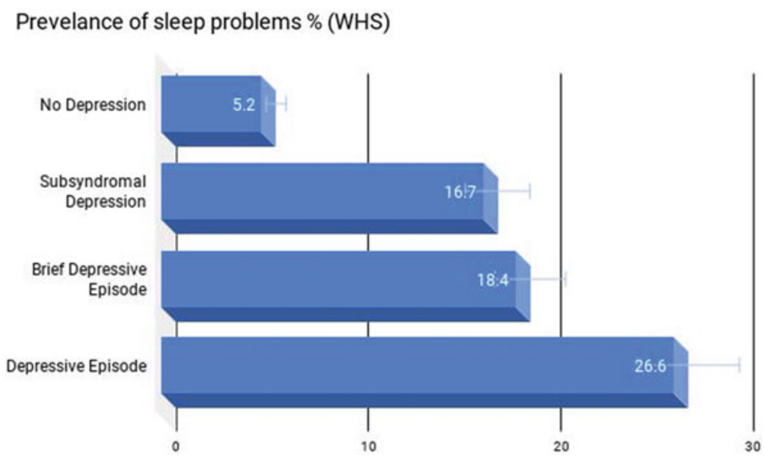



**Fig. 2** Prevalence of sleep problems by different types of depression (95 percent confidence level) [25]

symptoms of depression were lower in physically active people [26]. A cumulative analysis is an approach that combines all these factors to provide greater accuracy (Fig. 3).

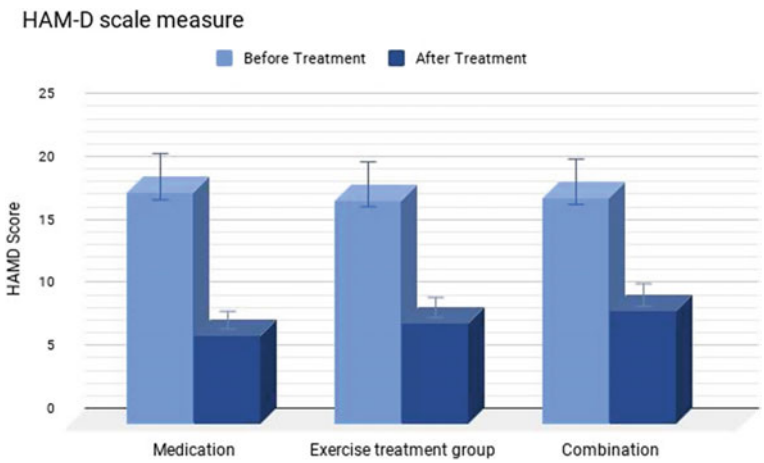



**Fig. 3** Observed mean Hamilton Rating Scale for depression before and after treatment [27]

### 3.1 Cumulative Analysis

1. Crowd-sourcing of data and mental health survey: Twitter data, details of sleep patterns, physical activities, and caloric intake was obtained from the control group. A survey is then conducted on these people in order to determine a diagnosis of depression in the test subjects. The obtained values are then used to correlate and verify the results derived from the classification algorithm.
2. Data from micro-blogging platforms: Crowd-sourcing is performed and text data is obtained from a group of people. This can be done from any social media platform where the user expresses his or her feelings using words.
3. Data of sleep patterns: Sleep cycle and quality of sleep are measured in the people that are a part of the crowd-sourcing. This data is obtained using simple sensors in one's mobile, these sensors determine the movement during sleep, using this the quality of sleep is measured.
4. Data of physical activities: Physical activities of the individuals participating in the crowd-sourcing is obtained. Low power location services and the motion sensors on one's phone detect activity types for outdoor activity and measure the intensity of an indoor activity. Based on that minutes moved by an individual is calculated and calories burnt throughout the day are estimated.
5. Data of caloric intake: The caloric values of each meal consumed by the subjects were recorded and aggregated in order to ascertain the total caloric intake. The average caloric intake was then derived from the total calories consumed by the subject group over the period of observation. This data can be recorded and submitted by the subject through self-documentation.
6. Classification algorithm: The sleep patterns obtained from the sensors of people participating in crowd-sourcing are compared with the sleep patterns of people that are not going through any mental illness in several different age group categories. An algorithm is deployed to map this variation for a prolonged period since a variation caused just for a few days may be due to other external factors. The same method is used to determine deviation in the physical activities and caloric intake of an individual.
7. Cumulative analysis result: The result is derived from the sentiment analysis of the twitter data in conjunction with the physiological data obtained through crowd-sourcing. The classification algorithm is then trained on this amalgamated data (social media activities, sleep, caloric intake, and physical activities) to produce a prediction of depression diagnosis in the subject. Since the result obtained factors in a lot of other features, it produces a higher accuracy (Fig. 4).

### 3.2 Mathematical Model

The approach taken presents a system that identifies users at the risk of depression and is done on the basis of the text comprising their tweets and their associated physiological data such as food habits (caloric intake per day) and sleep pattern

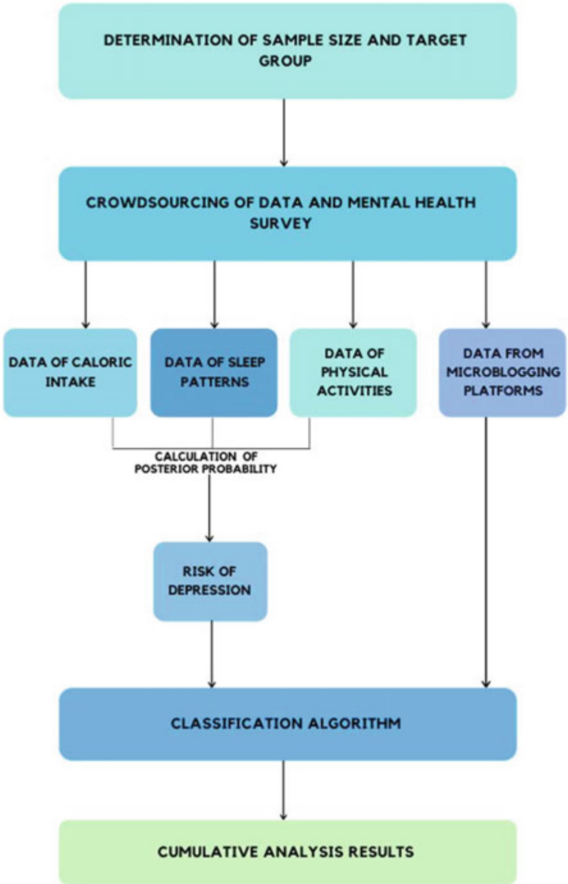



**Fig. 4** Flowchart of proposed cumulative analysis

(hours of sleep per day). The algorithm used for this is based on the Multinomial Naive Bayes method.

A Naive Bayes classifier is a fairly simple, yet accurate mathematical model that is based on the Bayes Theorem of Probability in statistics. It is fairly on par in terms of accuracy and precision with classification algorithms such as Support Vector Machines, Logistic Regression, and others [28]. It operates under certain presuppositions such as each feature being independent of another, in essence, probabilities of occurrence of each event (in this case, words) are mutually exclusive and have no relation to the other. This greatly simplifies the computational complexity of the algorithm for a given body of the text.

Under the independence assumption of the Naive Bayes algorithm, each word in the tweet that is being analyzed is treated independently. The assumption is that the effect of the value of a predictor($x$) on a given class ($c$) is independent of the values of other predictors, which is known as class conditional independence. Equation (1) shows the calculation of the value of posterior probability $P(c|x)$ from $P(c)$, $P(x)$, and P(x|c), where $P(c|x)$ is the posterior probability of class (target) given predictor (attribute), $P(c)$ is the prior probability of class, $P(x|c)$ is the likelihood which is the probability of predictor given class and $P(x)$ is the prior probability of predictor [29].

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{1}$$

Equations (2) and (3) detail the Naive Bayes algorithm [16], while Eq. (4) details the working equation of the Multinomial approach.

$$P(c_k|x) = \frac{P(x|c_k)P(c_k)}{P(x)} \tag{2}$$

$$P(c_k|x_1, \ldots, x_n) = \frac{1}{Z}P(c_k)\prod_{i=1}^{n} P(x_i|c_k) \tag{3}$$

The Multinomial approach of the Naive Bayes method lends to have a higher degree of accuracy resulting from the greater number of variables or factors being considered. An increase in the number of variables in a multinomial distribution leads to an increase in the bias and decrease in the variance of the model used.

$$(x|c_k) = \frac{\left(\sum_i x_i\right)!}{\prod_i x_i!}\prod_i P_{k_i}^{x_i} \tag{4}$$

The posterior probability of a diagnosis of depression in a subject based on his or her sleep pattern is calculated by taking the standard deviation from the standard average sleep cycle for a healthy human being of 7–9 h [30] and dividing it by the average number of hours of sleep of the entire test group. The calculation of the probability of depression based on food habits is also conducted in a similar manner with the standard caloric intake being set at 2000–2500 cal per day [31]. The posterior probability of depression based on physical exercise is calculated with the standard exercise time for an adult being estimated at 150–300 min of moderate aerobic exercise per week (caloric depletion of 260–600 cal per 30 min) [32].

## 4 Conclusion

The paper thoroughly reviews popular methods of sentiment analysis that are performed on social media data to determine depression amongst users based on their social media posts. The paper also suggests a more holistic method of performing sentiment analysis on social media that factors in additional data primarily of which are sleep patterns and physical activities. This additional information allows us to obtain a more accurate measure of depressed users thus reducing the probability of misclassification. It also mitigates the effects of misclassification due to the inability to measure context or humor in social media posts, allowing for more useful data to be derived.

One of the implications of performing this method is that the data that is required to make the determination is not available for the general users on social media. The data has to be voluntarily documented and submitted by the user in order to conduct analysis through this method. This negates the ability to perform large scale analysis on social media platforms.

Another effect of using this method is that the results obtained are skewed towards the geographic area in which the user resides. This can heavily affect the extent to which the classification is aligned towards either end of the sentiment spectrum.

Further improvement in the accuracy of this method can be done by taking into consideration more factors such as dietary habits comprising a more detailed report of the dietary composition, in-depth recorded observations of the sleep patterns, and pre-existing health conditions. This would help in achieving a more accurate classification result of the user depression metrics as it improves the multivariate analysis result obtained.

## References


1. World Health Organization (2019)\Depression. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression
2. De Choudhury MM, Gamon M, Counts S, Horvitz E (2013) Predicting depression via social media
3. Dea BO, Wan S, Batterham PJ, Calear AL, Paris C, Christensen H (2015) Detecting suicidality on twitter. Internet Interv 2(2):183–188. https://doi.org/10.1016/j.invent.2015.03.005
4. Ayers JW, Althouse BM, Dredze M (2014) Could behavioral medicine lead the web data revolution? JAMA—J Am Med Assoc 311(14):1399–1400. https://doi.org/10.1001/jama.2014.1505
5. Wang W, Chen L, Thirunarayan K, Sheth AP (2012) Harnessing twitter big data for automatic emotion identification. In: 2012 proceedings of the ASE/IEEE international conference on privacy, security, risk and trust 2012 and ase/ieee international conference social computing, pp 587–592, 2012. https://doi.org/10.1109/SocialCom-PASSAT.2012.119
6. Pang LLB (2008) Opinion mining and sentiment analysis
7. Jáuregui-Lobera I (2011) Neuroimaging in eating disorders. Neuropsychiatr Dis Treat 7:577–584. https://doi.org/10.2147/NDT.S25186
8. Varady KA, Hellerstein MK (2007) Alternate-day fasting and chronic disease prevention: a review of human and animal trials. Am J Clin Nutr 86(1):7–13
9. Chacón F, Esquifino AI, Perelló M, Cardinali DP, Spinedi E, Alvarez MP (2005) 24-hour changes in ACTH, corticosterone, growth hormone, and leptin levels in young male rats subjected to calorie restriction. Chronobiol Int 22(2):253–265. https://doi.org/10.1081/CBI-200053522
10. Oxman TE, Rosenberg SD, Tucker GJ (1982) The language of paranoia. Am. J. Psychiatry
11. De Choudhury M, Counts S, Horvitz E (2013) \Social media as a measurement tool of depression in populations. In: Proceedings of the 5th annual ACM web science conference, 2013
12. Rude SS, Valdez CR, Odom S, Ebrahimi A (2003) Negative cognitive biases predict subsequent depression. Cogn Ther Res. https://doi.org/10.1023/A:1025472413805

13. G. Coppersmith, M. Dredze, and C. Harman, "Quantifying Mental Health Signals in Twitter," pp. 51–60, 2015, doi: https://doi.org/10.3115/v1/w14-3207.
14. Rude SS, Gortner EM, Pennebaker JW (2004) Language use of depressed and depression-vulnerable college students. Cogn Emot. https://doi.org/10.1080/02699930441000030
15. "LIWC | Linguistic Inquiry and Word Count." http://liwc.wpengine.com/ (accessed Jun. 27, 2020).
16. Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC (2017) Detecting depression and mental illness on social media: an integrative review. Curr Opin Behav Sci 18:43–49. https://doi.org/10.1016/j.cobeha.2017.07.005
17. Nadeem M (2016) Identifying Depression on Twitter, pp 1–9 [Online]. Available: http://arxiv.org/abs/1607.07384.pdf
18. Hamilton M (1989) Frequency of symptoms in melancholia (depressive illness). Br J Psychiatry 154(2):201–206
19. Yates WR, Mitchell J, John RA et al (2007) Clinical features of depression in outpatients with and without co-occurring general medical conditions in STAR*D: confirmatory analysis. Prim Care Companion J Clin Psychiatry 9:7–15
20. Paterson LM, Nutt DJ, Wilson SJ (2008) NAPSAQ 1: national patient sleep assessment questionnaire in depression. Int J Psychiatry Clin Pract (in press)
21. Mayers AG, Van Hooff JC, Baldwin DS (2003) Quantifying subjective assessment of sleep and life-quality in antidepressant-treated depressed patients. Hum Psychopharmacol 18:21–27
22. Nutt DJ, Wilson S, Paterson L (2008) Sleep disorders as core symptoms of depression. Dialogues Clin Neurosci
23. Benca RM, Obermeyer WH, Thisted RA, Gillin JC (1992) Sleep and psychiatric disorders: a meta-analysis. Arch Gen Psych 49:651–670
24. Maquet P, Degueldre C, Delfiore G et al (1997) Functional neuroanatomy of human slow-wave sleep. J Neurosci 17:2807–2812
25. Drevets WC (2007) Orbitofrontal cortex function and structure in depression. Ann N Y Acad Sci 1121:499–527
26. Stickley A, Leinsalu M, DeVylder JE, Inoue Y, Koyanagi A (2019) Sleep problems and depression among 237 023 community-dwelling adults in 46 low-and middle-income countries. Sci Rep. https://doi.org/10.1038/s41598-019-48334-7
27. Kleppang AL, Hartz I,Thurston M, Hagquist C (2018) The association between physical activity and symptoms of depression in different contexts—A cross-sectional study of Norwegian adolescents. BMC Public Health 18(1). https://doi.org/10.1186/s12889-018-6257-0
28. Blumenthal et al (1999) Effects of exercise training on older patients with major depression. Arch Intern Med. https://doi.org/10.1001/archinte.159.19.2349
29. Rish I (2001) An empirical study of the naive Bayes classifier. IJCAI 2001 Work Empir methods Artif Intell 22230:41–46
30. Naive Bayesian: https://www.saedsayad.com/naive_bayesian.htm#:~:text=Naive%20Bayes%20classifier%20assume%20that,)%20given%20predictor%20(attribute)
31. How much sleep do we really need?—Harvard Health. https://www.health.harvard.edu/staying-healthy/how-much-sleep-do-we-really-need
32. What should my daily intake of calories be?—NHS. https://www.nhs.uk/common-health-questions/food-and-diet/what-should-my-daily-intake-of-calories-be/
33. Jindal RD, Thase ME (2004) Treatment of insomnia associated with clinical depression. Sleep Med Rev 8:19–30
34. Katz DA, McHomey CA (2002) The relationship between insomnia and health-related quality of life in patients with chronic illness. J Fam Pract 51:229–235
35. Calories burned in 30 minutes for people of three different weights—Harvard Health. https://www.health.harvard.edu/diet-and-weight-loss/calories-burned-in-30-minutes-of-leisure-and-routine-activities

# Process Logo: An Approach for Control-Flow Visualization of Information System Process in Process Mining

**M. V. Manoj Kumar, B. S. Prashanth, H. R. Sneha, Likewin Thomas, B. Annappa, and Y. V. S. Murthy**

## 1 Introduction

Process mining offers a range of techniques for extracting useful information from the event logs [1, 8]. At the outset, techniques in process mining can broadly be classified into three categories, i.e., discovery, conformance, and enhancement.

- Discovery [2] methods generate visual control flow models (similar to the node and edge graphs).
- Conformance [3] techniques aim at cross-validating the discovered control flow model with the actual information recorded in the event log.
- Enhancement [4] methods add additional information on the discovered model to make it more informative.

The focus of this paper is to examine the various techniques related to process discovery. Identifying and formulating the process model has been a predominant topic in process mining research. The most commonly used discovery techniques usually generate node-edge graphs (for example, Figs. 2 and 3 show control flow

M. V. Manoj Kumar (✉) · B. S. Prashanth · H. R. Sneha
Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, India

L. Thomas
Department of Computer Science and Engineering, PES Institute of Technology and Management, Shivamogga, India

B. Annappa
Department of Computer Science and Engineering, National Institute of Technology, Karnataka, Surathkal, India
e-mail: annappa@ieee.org

Y. V. S. Murthy
School of Computing, Vellore Institute of Technology, Vellore, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_40

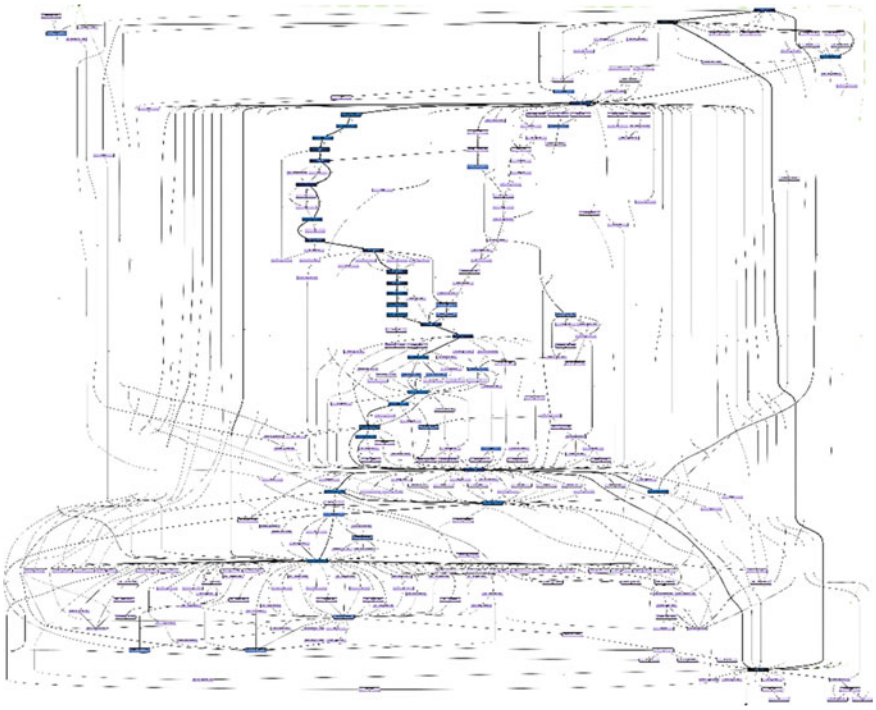


**Fig. 1** Basic techniques in process mining (discovery, conformance, and enhancement)

modeled in Petri net [2] and Transition net [5] modeling notation). In the case of large data sets, the discovered control flow model overflows with information rendering end-users bewildered. Also tracking the connection between several nodes becomes difficult if the number of edges increases over time. One such example is shown in Fig. 1. The control flow model is related to the Hospital Billing event log which comprises events related to billing of medical facilities.[1]

The "process logo" control flow visualization method demonstrated in this paper overcomes the drawback of existing methods, it generates comprehensible and simple control flow visualization irrespective of event log complexity.

The upcoming sections of this paper are organized as follows, Sect. 2 briefs the related work, structure of a typical event log, i.e., input data for constructing process logo is discussed in Sect. 3. Section 4 gives a detailed description of the framework proposed for extracting process logo from a sample event log. Section 5 briefs the mathematical background and methods used to realize the framework. Sections 6 and 7 discusses the results obtained from the experimentation and sheds insights on future research prospects.

[1] Event log of this hospital billing process can be downloaded from 4TU. Center for Research Data, can be downloaded from https://data.4tu.nl/repository/uuid:76c46b83-c930-4798-a1c9-4be94dfeb741.

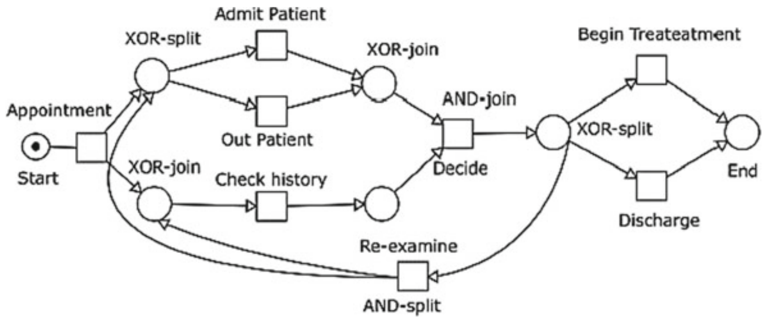



**Fig. 2** Control flow modeled in Petri net notation

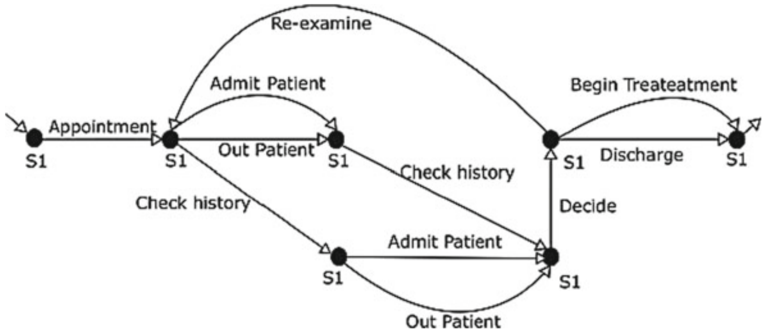



**Fig. 3** Control flow modeled in transition net notation

## 2 Related Work

Traditional Process Mining algorithms generate the control flow process model in terms of nodes connected by directed edges. In this approach, each node is an activity and a directed edge signifies the flow of the process from one activity to another activity. The following are some of the models generated by traditional control flow discovery methods in process mining. The control flow model is shown in Figs. 2 and 3 are related to the hospital treatment process, where, both the models convey similar meaning though they follow different modeling notation.

The following are the drawback of traditional methods for control flow visualization.

- It does not give the relative significance of the activities in the process (i.e., frequency of execution of an activity compared to other activities in the process)
- If the number of activities exceeds some limit, it becomes extremely difficult to draw viable conclusions from it.
- It lacks a visual description of the predominant path in the process.

The existing control flow methods can be improved to include this missing information visually, and make it more informative. By considering all these drawbacks, this paper proposes a new visualization method named "Process Logo" for comprehending the control flow perspective of the process more informatively.

# 3 Event Log

## *3.1 Structure of Event Logs*

Availability of an event log marks the starting point of process mining, the typical structure of an event log is shown in Fig. 4.

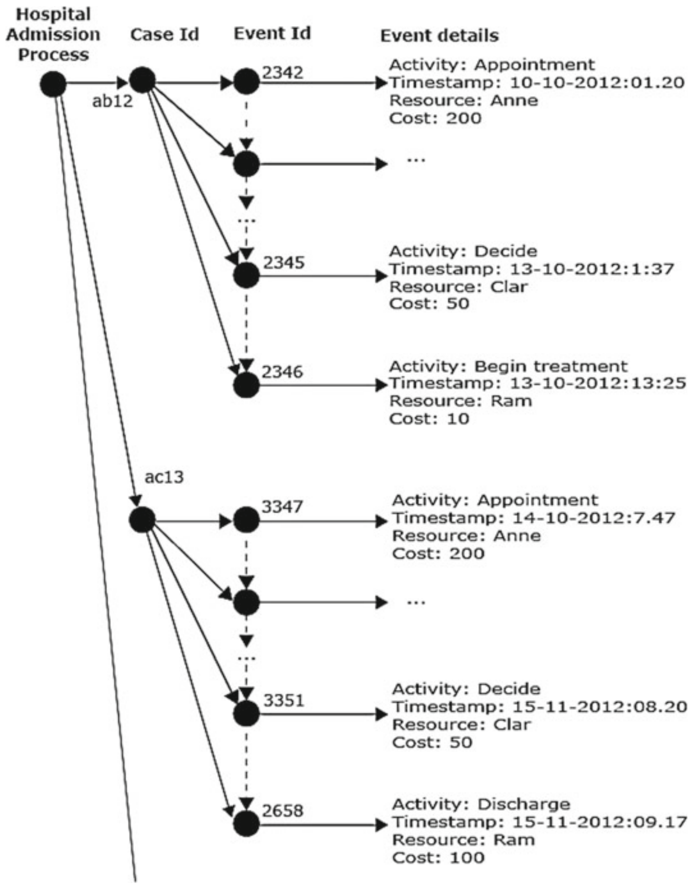



**Fig. 4** Structure of an event log (Data of hospital treatment process)

`ℒ =[⟨ Appointment, Admit patient, Check history, Decide, Begin treatment ⟩, ⟨ Appointment, Out patient, Check history, Decide, Discharge⟩, ⟨ Appointment, Admit patient, Check history, Decide, Re-examine, Out patient, Check history, Decide, Discharge ⟩, ⟨ Appointment, Check history, Admit patient, Decide, Discharge⟩, ⟨ Appointment, Out patient, Check history, Decide, Re-initiate, Admit patient, Check history, Decide, Begin treatment ⟩, ..]`

**Fig. 5** Control flow traces of hospital treatment process

Event logs are a collection or a bag of cases, with each case containing a sequence of events, each event in the case consists of activity name, timestamp, cost, resource, etc. The sequence of events in a case is collectively referred to as a trace. This paper only focuses on visualizing the control flow perspective (i.e., the sequence of execution of activities in a case and process altogether). The simplified version of the process log highlighting the control flow perspective is shown in Fig. 5. It represents a sequence of activities executed in cases related to the hospital treatment process.

For final model simplicity, each of the activities will be referred to by a one-letter equivalent and processed to extract the visual model.

## 4 Framework

The framework for the extracting process logo from an event log is given in Fig. 6. Construction of the process logo begins with the preprocessing step. In this step, all columns except activity columns are filtered out. Activity related to each case is collated and converted to a set of traces. A short name (typically a letter or an alphabet) is assigned to each activity for simplicity during the processing phase. For example, the first trace given in Fig. 5 is converted as follows.

<Appointment, Adimit patient, Check history, Decide,
Begin Treatment> =
<A, B, C, D, E>

Once, all the traces in the event log are converted to sequences, as shown above, the Levenshtein distance and Affinity Propagation Algorithm [6] method is applied to group similar instances into multiple clusters (in section). Each cluster in a clustering signifies a coherent set of traces. Followed to this, each cluster of traces is processed by applying dynamic programming techniques to obtain a similarity matrix. The progressive alignment method is applied to each similarity matrix and its corresponding cluster to generate the aligned sequences [7–9]. Finally, the sequences are

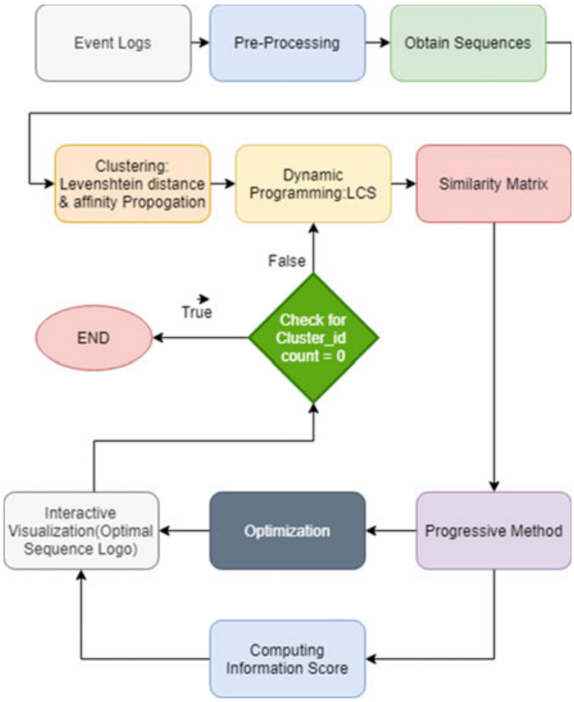



**Fig. 6** Framework for generating process logo from the event log

visualized using entropy, information score, consensus sequence, and data generated in the previous steps.

## 5 Levenshtein Distance and Affinity Propagation

With the help of clusters, we can provide information in a meaningful way in the case of large data sets specifically. For clustering sequences, we are using the Affinity Propagation Algorithm [6]. In this algorithm, the term *Exemplars* in the cluster are the points that describe other data points in the best possible way. The messages kept in,

- *Responsibility Matrix:* In this matrix *R*, *r(i,k)* determines how well point *k* can be considered as the exemplar for point *i*.
- *Availability Matrix:* In this matrix *A*, *a(i,k)* determines how well is point *i* to choose point *k* as its exemplar.

An example of a graph of 5 data points can be described in Fig. 7. A matrix can be composed of every connection *i* and *k* of the graph between two points.

In Fig. 7, the data points communicate between themselves and choose an exemplar (based on the distance between them). The exemplar is the most significant data

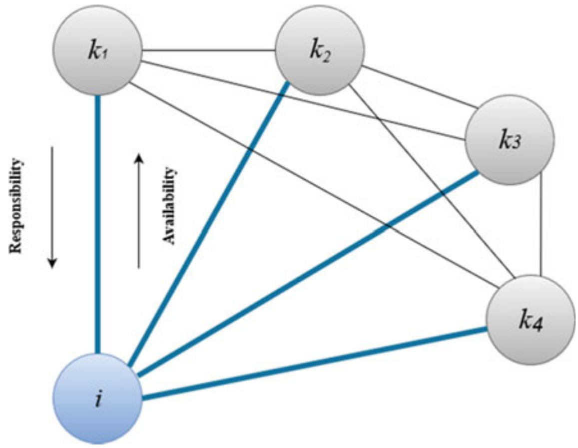



**Fig. 7** Graph of five data points

point of that cluster and the rest of the data points are part of that cluster. Responsibility and Availability are two matrices that define the clusters. Following is the algorithm to cluster the sequences into a similar group.

• The unaligned sequences are converted to an array to index it with a list.

• For every word split by in array words, the inverse of Levenshtein distance for $w_1$ in words to $w_2$ in words and store as similarity matrix $S$.

• For each row $i$, determine $max(A + S)$ of that row for every index not equal to $k$ or $i$. Where k is a function node (a possible exemplar) and i variable node. This is the responsibility matrix.

$$r(i, k) = s(i, k) - a(i, k') + s(i, k') \tag{1}$$

• The Availability matrix is determined by the responsibility at point $k$ and the sum of responsibilities that other data points assigned to $k$ for all points, not on the diagonal of $A$.

$$a(i, k) = (0, r(k, k)) + \sum max(0, r(i', k)) \tag{2}$$

$$a(k, k) = \sum max(0, r(i', k)) \tag{3}$$

• The maximum value of $A + B$ is used for selecting the final exemplar

$$exemplar(i, k) = max\{a(i', k) + b(i', k)\} \tag{4}$$

And when we apply the above algorithm to the set of sequences defined in the Results and Experiments section, we obtain the Message Graph as,

$$\mathbf{MessageGraph} = \begin{bmatrix} 0 & -5 & -5 & -4 & -6 & -5 \\ -5 & 0 & -4 & -2 & -9 & -2 \\ -5 & -3 & 0 & -2 & -9 & -2 \\ -4 & -2 & -2 & 0 & -9 & -3 \\ -6 & -9 & -9 & -9 & 0 & -9 \\ -5 & -2 & -2 & -3 & -9 & -0 \end{bmatrix}$$

The message graph matrix is considered as a similarity matrix, which is then used for aligning the sequences dynamically.

## 6 Results and Experiments

Suppose an example of the six sequences given below and cluster them using the above-defined algorithm. Consider the following six sequences for ease of understanding as shown in Table 1,

Here, each of the alphabets corresponds to an activity in the process. The following Table 2 summarizes the meaning attached to each alphabet for this example. Upon computing the Longest common subsequence we derive a table of similarity matrices

**Table 1** Example trace sequence

| Trace number | Trace |
|---|---|
| 1 | ABCDFCDG |
| 2 | ABCDECBDEBCDI |
| 3 | ABCDG |
| 4 | ACBDG |
| 5 | AFCDI |
| 6 | ACFDI |

**Table 2** Alphabets and its equivalent meaning

| Letter | Meaning |
|---|---|
| A | Register request |
| B | Examine casually |
| C | Check ticket |
| D | Decide |
| E | Reinitiate request |
| F | Examine thoroughly |
| G | Reject request |
| I | Pay compensation |

**Table 3** Matrix indicating the similarity between traces (sequences)

| Trace | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0 | 4 | 4 | 5 | 7 | 4 |
| 2 | 0 | 0 | 3 | 4 | 4 | 3 |
| 3 | 0 | 0 | 0 | 3 | 4 | 4 |
| 4 | 0 | 0 | 0 | 0 | 4 | 3 |
| 5 | 0 | 0 | 0 | 0 | 0 | 4 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 |

as given below in Table 3. Since the value of Trace 1 and Trace 5 results as the highest value in Table 3, we align Trace 1 and 5.

Later on, we align the resultant of the above alignment with Trace 4 as it is the next maximum value after the previously marked maximum value. By progressively working in a similar way, we get the Aligned Sequence as in Fig. 8.

These multiple aligned Traces can be leveraged to render the Optimized Process Logo which is represented in Fig. 9.

Figure 9, describes the visualization of the aligned sequences. Multiple aligned sequences from Fig. 8 are rendered into an optimized logo. This logo helps in understanding how the processes occur.

Example 2: Using clustering, the initial set of traces are clustered and then aligned.

Cluster 1:

- ABCDG (Exemplar)
- ABCDFCDG

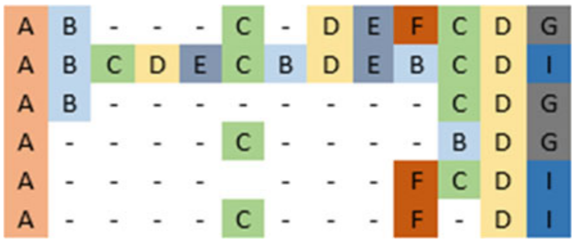



**Fig. 8** Sample of aligned traces

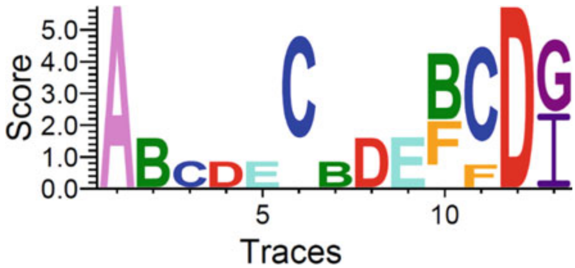



**Fig. 9** Sequence logo produced by multiple aligned sequences

**Fig. 10** Sequence logo for cluster 1

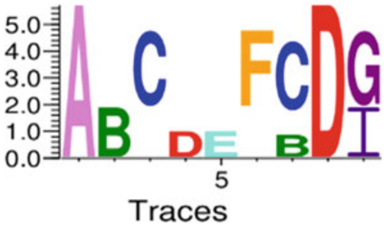
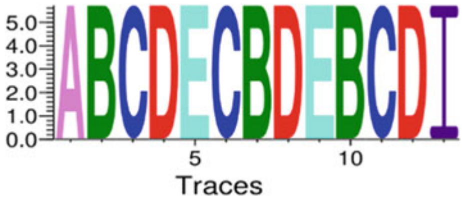



**Fig. 11** Sequence logo for cluster 2




- ACBDG
- AFCDI
- ACFDI

Cluster 2:

- ABCDECBDEBCDI (Exemplar)

The sequences marked as Exemplar are the Exemplar points chosen by cluster data points. After defining clusters, when we put this data to alignment and then produce a sequence logo, the logo is obtained as cluster 1 in Figs. 10 and 11.

In Fig. 10, with five sequences in cluster 1, the sequence logo forms as such. This is one part of the whole output, with other clusters forming such sequence logos.

In Fig. 12, interactive visualization of both the clusters, that is, Cluster 1 and Cluster 2 are merged into a single output. This helps in analyzing how the different clustered visualizations behave together.

## 7 Future Scope

Process mining has enormous scope in the imminent future. With IoT heading our way in a few year's time, mining a huge amount of events and processes will become tedious. Visualization of processes will require efficient techniques rather than the traditional way of Petri nets. Future development in this arena may result in visualizing the 3*D* models, where the results will be displayed as a part of PAISs [6] rendering it easy to understand the frequent processes.

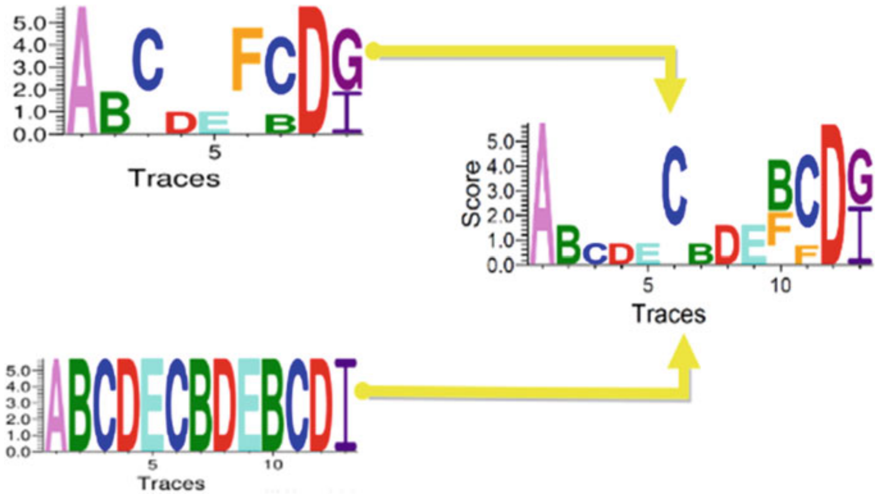



**Fig. 12** Creative visualization of a combination of both the clusters to obtain a single output

## 8 Conclusion

Traditional methods of control flow modeling and representation are done using directed graphs of nodes and edges(Petri Nets). These approaches are prone to become elaborated and tedious in representation with the increase in the number of activities involved in the Information System increase over time. The paper proposes a method called Process Logo, which provides a simplified representation of the control flow model which encompasses the representation of parameters such as the activity list, their order, information gain/score, major event, etc. Process logo wins over the traditional approaches in the amount of information involved that is visually represented for a large number of activities. The process logo is based on a biological sequence alignment technique called multiple-sequence alignment, applied to the sequences of the real-life processes. Though the proposed Process logo approach seems appropriate and efficient for control flow representation for a large number of activities in the information system, it does rely on the number of available alphabets and availability of Event logs.

## References

1. Van Der Aalst W (2012) Process mining. Commun. ACM 55(8):76–83
2. B.F. Van Dongen, A.A. De Medeiros, L. Wen, Process mining: overview and outlook of petri net discovery algorithms, in *Transactions on Petri Nets and Other Models of Concurrency II* (Springer, Berlin, Heidelberg, 2009), pp 225–242

3. Van der Aalst WM, de Medeiros AKA (2005) Process mining and security: detecting anomalous process executions and checking process conformance. Electron. Notes Theor. Comput. Sci. 121:3–21
4. W. Van Der Aalst, A. Adriansyah, A.K.A. De Medeiros, F. Arcieri, T. Baier, T. Blickle, A. Burattin, Process mining manifesto, in *International Conference on Business Process Management* (Springer, Berlin, Heidelberg, Aug 2011), pp 169–194
5. W.M. Van der Aalst, A.A. De Medeiros, A.J.M.M. Weijters, Genetic process mining, in *International Conference on Application and Theory of Petri Nets* (Springer, Berlin, Heidelberg, June 2005), pp 48–69
6. B.J. Frey, D. Dueck, Clustering by passing messages between data points. Science **315**(5814), 972–976 (2007)
7. Hunt JW, Szymanski TG (1977) A fast algorithm for computing longest common subsequence. Commun. ACM 20(5):350–353
8. Dumas M, Van der Aalst WM, Ter Hofstede AH (2005) Process-Aware Information Systems: Bridging People and Software Through Process Technology. Wiley
9. Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73(1):237–244

# On the Maximum N-degree Energy of Graphs

**G. B. Sophia Shalini, B. V. Dhananjayamurthy, and Anwar Saleh**

## 1 Introduction

The theory of energy of graphs originated from chemistry, and mathematically, the definition of energy of a graph was given by I. Gutman [1] in the year 1978. Let $G$ be a graph with $|V(G)| = nand|E(G)| = m$. Let $A(G)$ represent the adjacency matrix of $G$ and $\lambda_1, \lambda_2, \ldots \lambda_n$ be the n eigenvalues of $A(G)$. Then, the energy of graph $E(G)$ is the sum of the absolute values of the eigenvalues of $A(G)$ of the graph $G$. That is,

$$E(G) = \sum_{i=1}^{n} |\lambda_i| \tag{1}$$

For all comprehensive work related to energy of graphs, read [2–5] with the references mentioned therein. The elementary results together with numerous lower and upper bounds of graph energy have been developed in [6, 7], and for the chemical applications of energy of graphs, refer the papers [8, 9].

Chandrashekar Adiga and M. Smitha [10] introduced the concept of maximum degree energy of a graph. Motivated by the recent works on the energy of graphs, we have introduced the concept of maximum $N$-degree energy of graphs.

G. B. Sophia Shalini
Department of Mathematics, Mount Carmel College, Bangalore-58, India

B. V. Dhananjayamurthy (✉)
Department of Mathematics, Nitte Meenakshi Institute of Technology, Yelahanka, Bangalore-64, India
e-mail: Dhananjayamurthy.bv@nmit.ac.in

A. Saleh
Department of Mathematics, College of Science, University of Jeddah, Jeddah, Saudi Arabia
e-mail: Asaleh@uj.edu.sa



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_41

Consider a simple connected graph $G$ with the vertices $v_1, v_2, \ldots, v_n$. Let us define the $N$-degree of a vertex $v_i$ as the sum of the degree of the vertices in the open neighborhood of $v_i$. That is,

$$N_d(v_i) = \sum_{u \in V(G)} d(u), u \in N(v_i)$$

Let $N_i$ be the $N$-degree of the vertex $v_i$. Define

$$N_{ij} = \begin{cases} \max\{N_i, N_j\} & if \;\; v_i v_j \in E(G) \\ 0 & \text{otherwise} \end{cases}$$

Then, the $n \times n$ matrix $N_d(G) = [N_{ij}]$ is referred as the maximum $N$-degree matrix of $G$, and the characteristic equation of maximum $N$-degree matrix is

$$f(G;\lambda) = \det(\lambda I - N_d(G)) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \ldots + c_n = 0$$

where $I$ is the $n$th-order identity matrix. The eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$ of $N_d(G)$ are real, since the $N$-degree matrix $N_d(G)$ is a real symmetric matrix and assuming them in a non-increasing order, say $\lambda_1 \geq \lambda_2 \geq \cdots\cdots \geq \lambda_n$. Then, the maximum $N$-degree energy of a graph is defined as follows:

$$E_N(G) = \sum_{i=1}^{n} |\lambda_i|$$

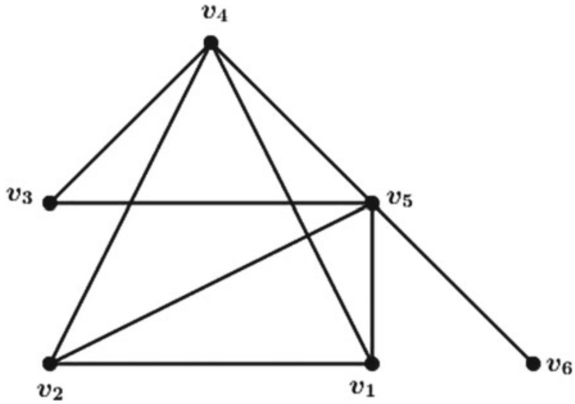
If the distinct maximum $N$-degree eigenvalues of $N_d(G)$ are $\lambda_1 > \lambda_2 > \cdots\cdots > \lambda_r, r \leq n$ with $m_1, m_2, \ldots\ldots, m_r$ as the multiplicities, then we shall write the spectrum of $G$ as

$$spec(G) = \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_r \\ m_1 & m_2 & \ldots & m_r \end{pmatrix}.$$

**Example 1.** Let $G_1$ be a graph in Fig. 1. The maximum $N$-degree matrix of $G_1$ is given by.

$$N_d(G_1) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 12 \\ 0 & 0 & 13 & 13 & 13 & 13 \\ 0 & 13 & 0 & 0 & 0 & 12 \\ 0 & 13 & 0 & 0 & 12 & 12 \\ 0 & 13 & 0 & 12 & 0 & 12 \\ 12 & 13 & 12 & 12 & 12 & 0 \end{pmatrix}$$

**Fig. 1** Graph $G_1$




The characteristic polynomial of $N_d(G_1)$ is $f(G_1, \lambda) = \lambda^6 - 1396\lambda^4 - 19680\lambda^3 + 65808\lambda^2$, and the maximum $N$-degree eigenvalues are $-22.6207$, $-16.8516$, $-12.0000$, $3.0590$, $5.8890$, $42.5239$.

## 2 Bounds for Maximum *N*-degree Energy and Maximum *N*-degree Eigenvalues

In this section, we obtain the values of some coefficients of the characteristic equation of maximum $N$-degree matrix $N_d(G)$. It is obvious that $c_0 = (-1)^n$ and $c_1 = traceN_d(G) = 0$.

(i) We have $c_2 = \sum_{1\le j<,k\le n} \begin{vmatrix} N_{jj} & N_{jk} \\ N_{kj} & N_{kk} \end{vmatrix} = \sum_{1\le j<k\le n} \begin{vmatrix} 0 & N_{jk} \\ N_{kj} & 0 \end{vmatrix}$

But $\begin{vmatrix} 0 & N_{jk} \\ N_{kj} & 0 \end{vmatrix} = \begin{cases} -\{\max(N_j, N_k)\}^2, & if\ v_j \text{ and } v_k \text{ are adjacent} \\ 0, & \text{otherwise} \end{cases}$

Thus, $c_2 = -\sum_{i=1}^{n}(a_i + b_i)N_i^2$, where $a_i$ is the number of vertices in the neighborhood of $v_i$ whose $N$-degrees are less than $N(v_i)$ and $b_i$ is the number of vertices $v_j, (j < i)$ in the neighborhood of $v_i$ whose $N$-degrees are equal to $N(v_i)$.

It is easy to observe that the size of the graph $G = \sum_{i=1}^{n}(a_i + b_i)$.

**Example 2.** Consider the graph $G_1$ in Fig. 1. The coefficient $c_2$ of $\lambda^4$ in the characteristic equation of $N_d(G_1) = -\sum_{i=1}^{n}(a_i + b_i)N_i^2$

$$= -\left[(0+0)5^2 + (4+0)3^2 + (0+0)9^2 + (0+2)12^2 + (0+1)12^2 + (2+0)12^2\right]$$

$$= -[676 + 288 + 144 + 288] = -1396.$$

(ii) We have $c_3 = (-1)^3 \sum_{1 \le i < j < k \le n} \begin{vmatrix} N_{ii} & N_{ij} & N_{ik} \\ N_{ji} & N_{jj} & N_{jk} \\ N_{ki} & N_{kj} & N_{kk} \end{vmatrix}$

$$= -2 \sum_{1 \le i < j < k \le n} N_{ij} N_{jk} N_{ki} = -2 \sum_{\substack{\triangle\ v_i v_j v_k \\ N(v_i) \le N(v_j) \le N(v_k)}} [N^2(v_k) N(v_j)].$$

It is to be noted that the number of terms in the above sum is same as the number of triangles in the given graph. Also, if there is no triangle in $G$, then $c_3 = 0$.

**Example 3.** The coefficient $c_3$ of $\lambda^3$ in $N_d(G_1)$ is

$$= -2 \sum_{\substack{\triangle\ v_i v_j v_k \\ N(v_i) \le N(v_j) \le N(v_k)}} N^2(v_k) N(v_j)$$

$$= -2\left[13^2.12 + 13^2.12 + 13^2.12 + 12^2.12 + 13^2.12\right]$$

$$= -19680.$$

**Theorem 1.** If $\lambda_1, \lambda_2, \ldots..\lambda_n$ are the eigenvalues of the maximum $N$-degree matrix of the graph $G$, then $\sum_{i=1}^{n} \lambda_i^2 = -2c_2$.

Proof. We have

$$\sum_{i=1}^{n} \lambda_i^2 = \text{trace} N_d^2(G) = \sum_{i=1}^{n} \left( \sum_{k=1}^{n} N_{ik} N_{ki} \right)$$
$$= 2 \sum_{i=1}^{n} (a_i + b_i) N^2(v_i) = -2c_2$$

## 3 The Maximum $N$-degree Energy of Certain Families of Graphs

**Theorem 2.** If $\boldsymbol{K_n}$ is the complete graph on n vertices, then $\boldsymbol{E_N(K_n)} = 2(\boldsymbol{n} - 1)^3$.

Proof. We have

$$|\lambda I - N_d(K_n)| = \begin{vmatrix} \lambda & -(n-1)^2 & -(n-1)^2 & \cdots & -(n-1)^2 \\ -(n-1)^2 & \lambda & -(n-1)^2 & \cdots & -(n-1)^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -(n-1)^2 & -(n-1)^2 & -(n-1)^2 & \ldots & \lambda \end{vmatrix}$$

The characteristic polynomial is given by $\left(\lambda + (n-1)^2\right)^{n-1}\left(\lambda - (n-1)^3\right)$

$$\operatorname{spec}(K_n) = \begin{pmatrix} -(n-1)^2 & -(n-1)^3 \\ n-1 & 1 \end{pmatrix}$$

Therefore, $E_N(K_n) = \left|-(n-1)^2\right|(n-1) + \left|(n-1)^3\right| = 2(n-1)^3$.

**Theorem 3.** If $S_n$ is a star graph, then $E_N(S_n) = 2\sqrt{(n-1)^3}$.
Proof. We have

$$|\lambda I - N_d(S_n)| = \begin{vmatrix} \lambda & -(n-1) & -(n-1) & \cdots & -(n-1) \\ -(n-1) & \lambda & -(n-1) & \cdots & -(n-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -(n-1) & -(n-1) & -(n-1) & \ldots & \lambda \end{vmatrix}$$

The characteristic polynomial is given by $(-1)^n\lambda^{n-2}\left(\lambda + \sqrt{(n-1)^3}\right)\left(\lambda - \sqrt{(n-1)^3}\right)$

$$\operatorname{spec}(S_n) = \begin{pmatrix} 0 & -\sqrt{(n-1)^3} & \sqrt{(n-1)^3} \\ n-2 & 1 & 1 \end{pmatrix}$$

Therefore, $E_N(S_n) = |0|(n-2) + \left|-\sqrt{(n-1)^3}\right| + \left|\sqrt{(n-1)^3}\right| = 2\sqrt{(n-1)^3}$.

**Theorem 4.** If $K_{r,s}$ is a c complete bipartite graph, then $E_N\left(K_{r,s}\right) = 2\sqrt{(rs)^3}$.
Proof. We have

$$N_d\left(K_{r,s}\right) = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

where $A = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{r\times r}$ $,B = \begin{pmatrix} rs & rs & \cdots & rs \\ rs & rs & \ldots & rs \\ \vdots & \vdots & \ddots & \vdots \\ rs & rs & \cdots & rs \end{pmatrix}_{r\times s}$ and $C = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{s\times s}$

The characteristic polynomial is given by $(-1)^{r+s}\lambda^{r+s}\left(\lambda^2 - (rs)^3\right)$

$$\operatorname{spec}\left(K_{r,s}\right) = \begin{pmatrix} 0 & -(rs)^3 & (rs)^3 \\ r+s & 1 & 1 \end{pmatrix}$$

Therefore, $E_N\left(K_{r,s}\right) = |0|(r+s) + \left|-(rs)^3\right| + \left|(rs)^3\right| = 2\sqrt{(rs)^3}$.

**Theorem 5.** If $CP_n$ denotes the cocktail party graph $K_{r\times 2}$ on $n$ vertices, then $E_N(CP_n) = 4(n-1)(n-2)^2$.

Proof. Proceeding in the same way as above, the characteristic polynomial of cocktail party graph is given by $\left(\lambda + 2(n-2)^2\right)^{n-1}\lambda^n\left(\lambda - 2(n-1)(n-2)^2\right)$

$$\begin{aligned}
&\operatorname{spec}(CP_n) = \begin{pmatrix} -2(n-2)^2 & 0 & 2(n-1)(n-2)^2 \\ n-1 & n & 1 \end{pmatrix} \\
&E_N(CP_n) = \left|2(n-2)^2\right|(n-1) + 2(n-1)(n-2)^2 \\
&= 4(n-1)(n-2)^2
\end{aligned}$$

**Theorem 6.** If $S_n^0$ denotes the crown graph on $n$ vertices, then $E_N\left(S_n^0\right) = 4\left(\frac{n}{2}-1\right)^3$.
Proof. The characteristic polynomial of crown graph is

$$\left(\lambda + \left(\frac{n}{2}-1\right)^3\right)\left(\lambda + \left(\frac{n}{2}-1\right)^2\right)^{\frac{n}{2}-1}\left(\lambda - \left(\frac{n}{2}-1\right)^2\right)^{\frac{n}{2}-1}\left(\lambda - \left(\frac{n}{2}-1\right)^3\right)$$

$$\operatorname{spec}\left(S_n^0\right) = \begin{pmatrix} -\left(\frac{n}{2}-1\right)^3 & -\left(\frac{n}{2}-1\right)^2 & \left(\frac{n}{2}-1\right)^2 & \left(\frac{n}{2}-1\right)^3 \\ 1 & \frac{n}{2}-1 & \frac{n}{2}-1 & 1 \end{pmatrix}$$

Therefore, $E_N\left(S_n^0\right) = 4\left(\frac{n}{2}-1\right)^3$.

## 4 Conclusion

In this paper, we have defined the new matrix called maximum $N$-degree matrix and introduced the concept of maximum $N$-degree energy of a graph. We have computed the maximum $N$-degree energy of some standard families of graphs. Some properties of the characteristic polynomial of the maximum $N$-degree matrix and maximum degree eigenvalues are also explored. The concept can further be extended to find the maximum $N$-degree energy of molecular graphs.

## References


1. Gutman I (1978) The energy of a graph. Ber. Math-Statist. Sekt. Forschungsz. Graz. 103:1–22
2. D. Cvetkovic, I. Gutman, Applications of graph spectra, in Mathematical Institution. (Belgrade, 2009)
3. D. Cvetkovic, I. Gutman, Selected Topics on Applications of Graph Spectra, in Mathematical Institution, (Belgrade, 2011)
4. I. Gutman, The energy of a graph: Old and new results, in *Algebraic Combinations and Applications* eds. by A. Betten, A. Kohnert, A. Laue, A. Wassermann, (Springer, Berlin, 2001), pp 196–211. https://doi.org/10.1007/978-3-642-59448-9_13
5. Balakrishnan R (2004) The energy of a graph. Linear Algebra Appl. 387:287–295

6. H. Lui, M. Lu, F. Tian, Some upper bounds for the energy of graphs. J. Math. Chem. **41**(1), (2007)
7. McClelland BJ (1971) Properties of the latent roots of a matrix: The estimation of $\pi$-electron energies. J. Chem. Phy. 54:640–643
8. Graovac, I. Gutman, N. Trinajstic, Topological approach to the chemistry of conjugated molecules. (Springer, Berlin, 1977). https://doi.org/10.1007/978-3-642-93069-0
9. Gutman I, Polansky OE (1986) Mathematical Concepts in Organic Chemistry. Springer, Berlin
10. Adiga C, Smitha M (2009) On maximum degree energy of a graph. Int. J. Contemp. Math. Sci. 4(8):385–396

# Dynamic Resource Allocation for Virtual Machines in Cloud Data Center

**Niraj Kumar, Manan Kikla, and C. Navya**

## 1 Introduction

In the previous ten years, Data Centers (DCs) has developed in architectural way and not only in hardware resource and services. Many services such as storing data online to providing various types "apps" are made available to users. These resources make use of resources stored in cloud. The users are worried only about the quality, security and availability of resource and not where it is stored. All this is possible due to the role of (SDN) Software-Defined Networking paradigm leading in virtualization (Fig. 1).

## 2 Related Work

This section describes efforts made by several researchers to deal with issues and challenges in applying the methods for real-time applications. In [1], Raj Jain, Subharthi Paul show network virtualization is the key to the current and future success of cloud computing. In this article, we explain key reasons for virtualization and briefly explain several of the networking technologies that have been developed recently or are being developed in various standards bodies. In particular, we explain Software-Defined Networking, which is the key to network programmability. We also illustrate SDN's applicability with our own research on OpenADN—application delivery in a multi-cloud environment.

In [2], M. Gharbaoui B, Martini D, Adami G, Antichi S, Giordano, P. Castoldi show Oversubscription of intra-Data Center network links and high volatility of

N. Kumar · M. Kikla · C. Navya (✉)
Nitte Meenakshi Institute of Technology, Bangalore, India
e-mail: navya.c@nmit.ac.in



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_42

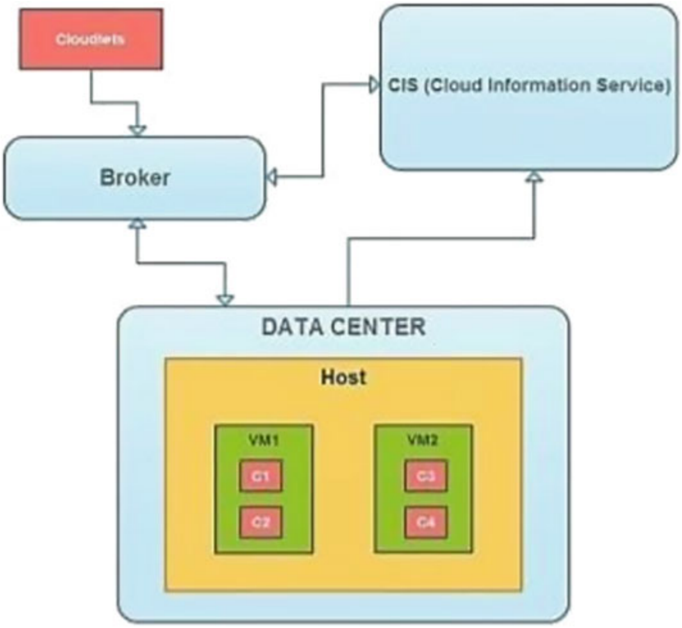



**Fig. 1** CloudSim model

VM deployments require a flexible and agile control of Data Center network infrastructures, also integrated with computing and storage resources. In this scenario, the Software-Defined Network paradigm and, specifically, the OpenFlow protocol, opens up new opportunities for the design of innovative resource management platforms that enable dynamic and fine-grain control of DC networks through traffic engineering algorithms. This paper investigates the performance of two different sets of cloud-fluent traffic engineering algorithms.

In [3], Adrian Lara, Anisha Kolasani, Byrav Ramamurthy show OpenFlow is currently the most commonly deployed Software-Defined Networking (SDN) technology. SDN consists of decoupling the control and data planes of a network. A software-based controller is responsible for managing the forwarding information of one or more switches; the hardware only handles the forwarding of traffic according to the rules set by the controller. OpenFlow is an SDN technology proposed to standardize the way that a controller communicates with network devices in an SDN architecture. It was proposed to enable researchers to test new ideas in a production environment. OpenFlow provides a specification to migrate the control logic from a switch into the controller. It also defines a protocol for the communication between the controller and the switches. As discussed in this survey paper, OpenFlow-based architectures have specific capabilities that can be exploited by researchers to experiment with new ideas and test novel applications.

# 3 Proposed System

Here, novel approach is proposed in which the clean separation between control and data planes is used which is a SDN key feature to take care of the critical issue in SDN framework. First, a comparison is made between the two versions of our allocators. Then a two-step allocation is done by disjoint allocator, in the first step location of the VM is chosen and then a proper network path is searched for. Computational and network requirements are taken care simultaneously in joint allocator, which is the advantage is the proposed system. Implement Genetic algorithm for resource allocation. This system can be used for comparing the joint and disjoint allocation strategies; in evaluating allocation strategies which are different it's either multi-objective or single objective and also used to optimize the network devices power consumption.

## *3.1 Implementation details*

The main goal is to accept as many virtual machines requests as possible and reduce network power consumption at the same time and this is done by of the IT Resource Allocator. Central Processing Unit, Random Access Memory, disk and bandwidth are the four parameters of a VM request.

The Steps for selection of server:

- Resource allocation to server
- Analytic ITRA
- Fuzzy IT Resource Allocator
- Genetic algorithm
- Power consumption and path computation

## *3.2 Resource allocation to server*

The most utilization of Central Processing Unit (CPU), Random Access Memory (RAM), disk and bandwidth are four resources are used to categories a VM request. We assign free CPU after the placement of the VM, free space in Random Access Memory (RAM), and after the placement of the VM, it is stored in the memory, and minimum cost path from servers to the external gateway (Fig. 2).

The availability of resources is taken care by the A-ITRA availability index and the computing for each candidate server the IA is done by analytic ITRA.

The combined version of the A-ITRA Availability Index is computed as follows:

$$I^s = 1/300[\text{CPUs} + \text{RAMs} + \text{DISKs}] + \alpha \text{PCs}/\text{PCM}$$

Only the computational resource is used to calculate the availability index when we use disjoint A-ITRA: (Fig. 3)

$$I^s A = 1/300[\text{CPUs} + \text{RAMs} + \text{DISKs}]$$

Fuzzy Logic Controller is designed with the help of fuzzy logic. The application of inference rules which are in the rule base is taken care by the inference engine is the core of the FLC. IF-THEN rules contain the system control strategies and are framed by premises and conclusions. Fuzzy rules use fuzzy sets and their associated membership to explain system variables is therefore fuzzification and defuzzification are two operations used for translations from conventional to fuzzy values. The mapping of input values into more fuzzy sets is fuzzification. While single conventional value to represent inferred values is done by defuzzification . Inputs of fuzzification are defined by four input fuzzy sets ("Not available", "Small", "Medium", and "Large") and its membrship functions. Figure 2 shows an example: while RAMs and DISKs are fuzzifiied in a single fuzzy set (i.e., "Medium" for RAM and "Large" for DISK), CPU with a certain probability from two different sets (0.7 in "Small" and 0.3 in "Medium");

The computation of gravity like the defuzzification algorithm for an aggregated fuzzy subset is done by defuzzification method ($f$).

$$\begin{aligned}
&I^s = f[\text{CPUs, RAMs, DISKs}] + \alpha\ \text{PCs/PCM} \\
&\text{Disjoint F} - \text{ITRA},\ I^s = f\ [\text{CPUs} + \text{RAMs} + \text{DISKs}] \\
&0 \le Disks \le 100\, 0 \le PCs \le PCM
\end{aligned}$$

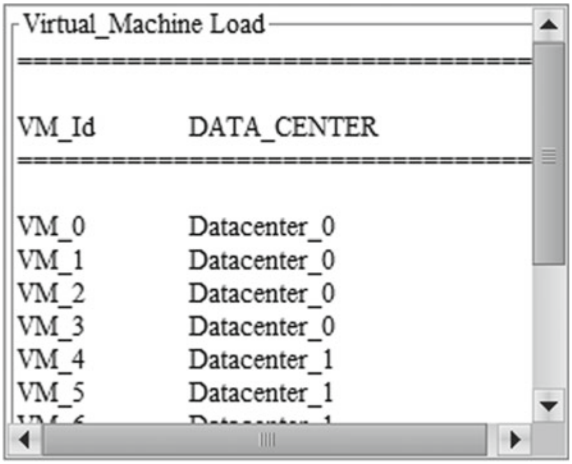



**Fig. 2** Resource allocation to server analytic ITRA

| DC_id | Hostid | BW | RAM | DISK(TB |
|---|---|---|---|---|
| Datacenter_0 | 0 | 25000 | 8192 | 5 |
| Datacenter_0 | 1 | 25000 | 16384 | 2 |
| Datacenter_1 | 0 | 25000 | 32768 | 1 |
| Datacenter_1 | 1 | 25000 | 16384 | 4 |

**Fig. 3** Analytic ITRA allocation fuzzy IT resource allocator

## *3.3 Genetic Algorithm*

In MODA allocation is done in multiple steps; first a list is created with the servers that match the request to allocate a new VM. Optimization defined between joint and disjoint MODA is the number of objectives (Fig. 4).

Disjoint MODA considers only CPUs, RAMs and DISKs, in this fist server selection is done then the minimum cost path is computed.

Joint MODA considers CPUs, RAMs, DISKs and PCs, during the sever selection phase the minimum cost path is associated.

Then, MODA starts by computing for the current location by working on the Pareto front, and then it makes a choice of possible allocations which follows applied policy (i.e., BF or WF) and one among the two possible strategies:

(1) Within the Pareto front randomly one solution is chosen (MODA-R);

Host Load

| VM_Id | DATA_CENTER |
|---|---|
| VM_0 | Datacenter_0 |
| VM_1 | Datacenter_0 |
| VM_2 | Datacenter_1 |
| VM_3 | Datacenter_1 |
| VM_4 | Datacenter_1 |
| VM_5 | Datacenter_1 |
| VM_6 | Datacenter_2 |
| VM_7 | Datacenter_2 |

**Fig. 4** Fuzzy ITRA allocation multi-objective dynamic allocator (MODA)

(2) Taking the solution which is at a min distance from the ideal vector and normalizing all the objectives (MODA-D).

The problem of minimization is solved by joint MODA as shown in formula when it obeys the BF policy:

$$\text{Minimize min }\{\text{CPUs};\ \text{RAMs};\ \text{DISKs};\ \text{PCs}\}\text{subject to}:\ 0 \leq \text{CPUs} \leq 100$$
$$0 \leq \text{RAMs0} \leq 100$$

The heuristics are strictly deterministic, it is done by considering the first virtual machine and deciding where it could be allocated, then the second virtual machine is considered and so on. By performing different permutations, the order of selection of virtual machine is changed this generates different solutions. We unnecessarily reduce the initial population size, to avoid inserting into the population an element that already is in it. After this initialization step, we do an iterative process, each cycle of it called a generation, where we:

1. Tournament process is used to choose two parents of a new individual.
2. Crossover operator is used to create a new individual.
3. Mutation operator is used to mutate some variables of new individual.
4. Keeping in consideration the unfeasibility due to overfilling we calculate the fitness of the individual.
5. Repair operator is used to fix the easy constraint.
6. The individual with lo west fitness is removed and an individual is inserted into the population.

### *3.4 Power consumption and path computation*

For switches, we consider power consumption model as follows the power model representing switches is:

$$P\text{switch} = P\text{chassis} + n \times P\text{line card} + P\text{load}$$

Where, Pswitch is the average power consumption of a single switch, Pchassis the constant power consumed by chassis, Pline card is the fixed amount of power consumed by line cards, Pload is power consumed (Fig. 5).

## 4 Use Case Diagrams

The figure given below is the use case diagram of proposed system, where the input is the dataset inputted by user, the algorithm works on the dataset to give the output.

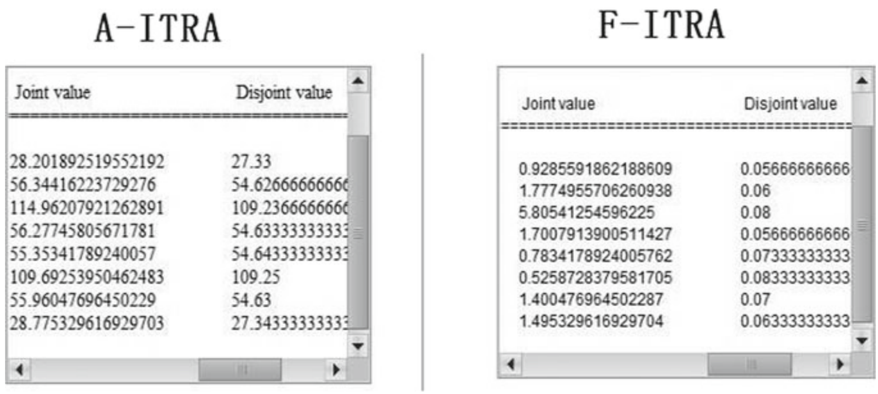



**Fig. 5** Power-aware allocation

An eclipse shape represents the use case namely VM request, resource computation and allocation problem (Fig. 6).

## 5 Results and Discussions

In the experiment conducted, different allocation strategies such as A-ITRA, F-ITRA are implemented considering two different parameters such as Joint and disjoint strategies. From the experiment conducted, it is identified that proposed schemes grouping all the VMs in minimum number of servers hence reducing the number of active computing devices (Figs. 7, 8, 9 and 10).

## 6 Conclusions

Power-aware dynamic virtual machines allocation in Cloud Data Centers takes advantage of cloud computing paradigm. Each VM request has four categories: Central Processing Unit, Random Access Memory, disk and bandwidth. Allocators are designed in order to take into consideration the power consumed by the network devices while accepting as many requests. In the proposed system, two different allocation strategies introduced are Analytic IT Resource allocation and Fuzzy IT Resource allocation for both joint and disjoint allocation. The experimental results are conducted to show two different allocation strategies of VM under data centers.

Power-aware dynamic virtual machine allocation has been taken into account with the resource requirements such as CPU, RAM, disk and bandwidth. We allocated all the network flows on the most power-efficient paths. In the proposed system we implemented joint and disjoint allocator in terms of no. of accepted requests. The

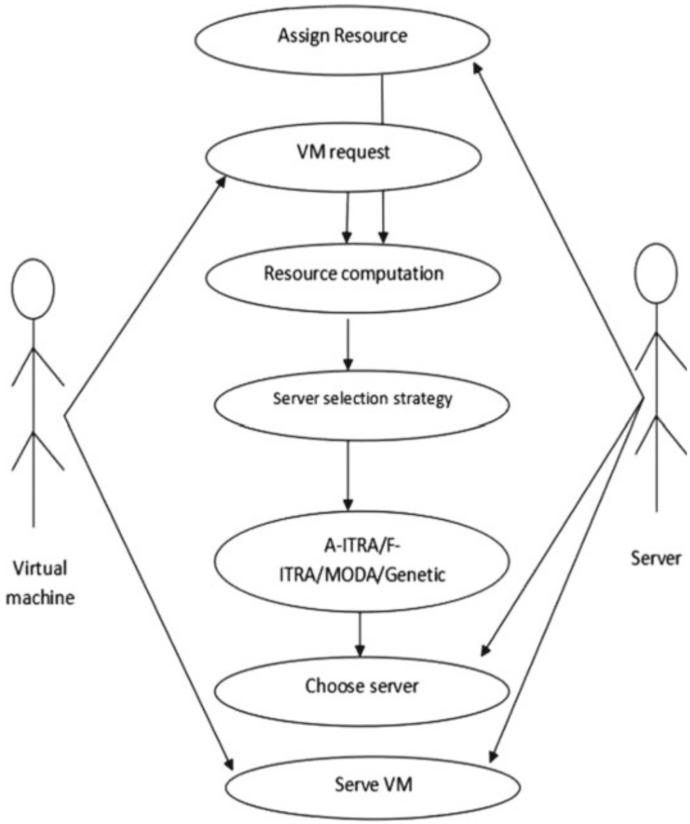



**Fig. 6** Use case diagram

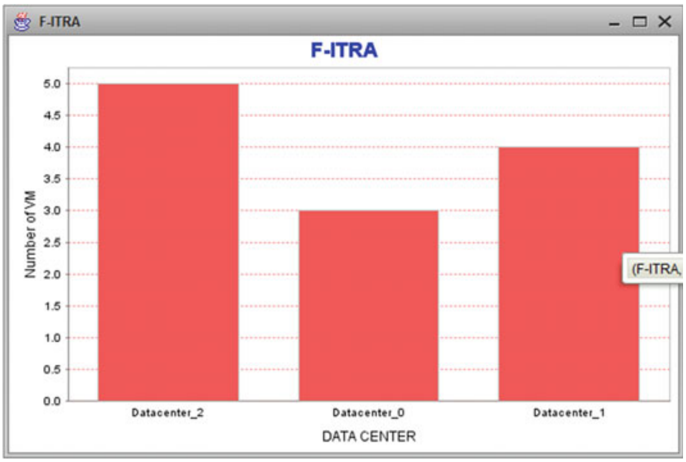



**Fig. 7** Allocation analytics using F-ITRA

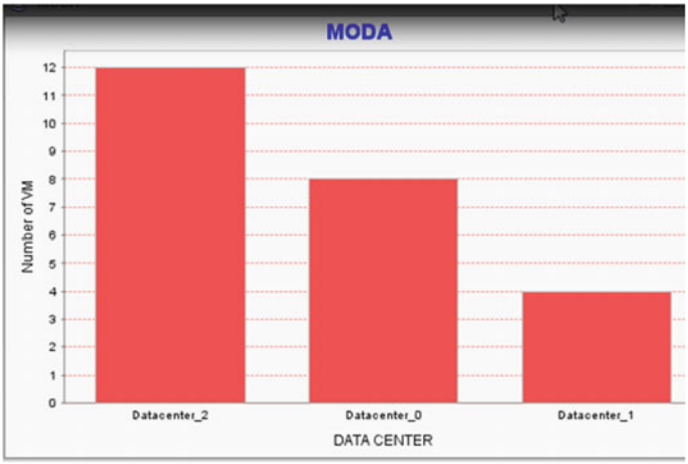



**Fig. 8** Allocation analytics using A-ITRA

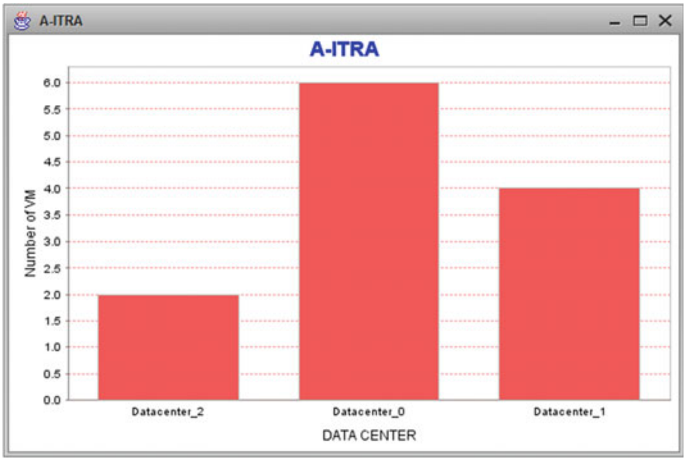



**Fig. 9** Allocation analytics using GA

two different strategies of virtual machine allocation is conducted namely F-ITRA and A-ITRA. The allocation is done and analyzed.

## 7 Future Works

Our work can possibly be extended in following directions:

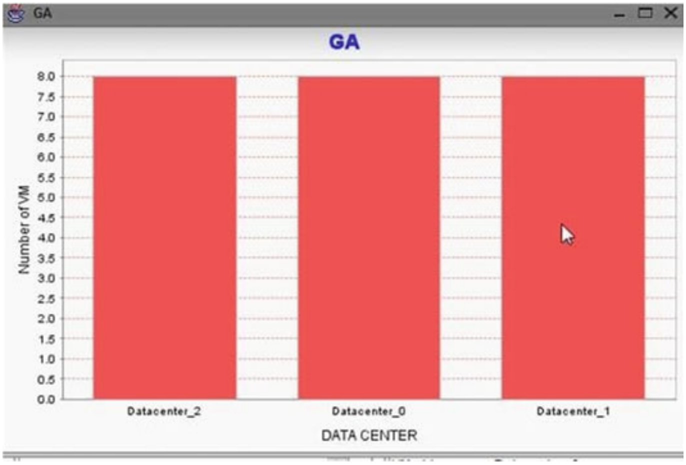



**Fig. 10** Allocation analytics using MODA

- For supporting live VM migration, trigger-based Immigration technique which is extended to include additional run time configuration management parameters.
- A commercial application, VM that is aware of energy consumption, reliable system with resource scheduling.
- With the help of deadline constrained scheduling algorithms, a fault tolerant scheduling and aware of energy consumption framework for cloud computing can be tested and developed.
- In heterogeneous environments, data offloading in cloud computing with the help of energy optimization algorithms.

## References

1. Jain R, Paul S (2013) Network virtualization and software defined networking for cloud computing: a survey. IEEE Commun. Magazine 51(11):24–31
2. M. Gharbaoui, B. Martini, D. Adami, G. Antichi, S. Giordano, P. Castoldi, On virtualization-aware traffic engineering in open flow data centers networks, in *2014 IEEE Network Operations and Management Symposium (NOMS)*. (IEEE, 2014), pp. 1–8. [Online]
3. Lara A, Kolasani A, Ramamurthy B (2014) Network innovation using open flow: a survey. IEEE Commun. Surv. Tutorials 16(1):493–512

# Image Captioning for the Visually Impaired

**Smriti P. Manay, Smruti A. Yaligar, Y. Thathva Sri Sai Reddy, and Nirmala J. Saunshimath**

## 1 Introduction

Image Captioning is defined as a process where a computer tries to perceive an image and give a textual representation for that image. For humans, it is quite simple and easy to describe a particular image due to a person's past experiences or maybe through some form of education. A quick glance at an image is sufficient for a human to point out and describe an immense amount of details about the visual scene [1]. For computers to do this it largely requires data with each image mapped to a set of captions. To achieve this, a technique called as Deep Learning is used. Deep Learning is defined as algorithms that provide Machine Learning more power to the existing Machine Learning algorithms to enhance its performance. Deep Neural Networks (DNNs) are extremely powerful machine learning models that achieve excellent performance on difficult problems such as speech recognition and visual object recognition [2]. It has the capability of handling missing data and also dealing with Terrabytes of data. Deep Learning works on the concept of how a human brain functions, i.e., using neurons. To apply the power of Deep Learning to a task like Image Captioning is a challenging task. There is a difficulty placing annotations on the correct regions [3]. Firstly, the suitable and relevant dataset must be available. As years of research has passed there are several datasets, namely Microsoft COCO dataset, the Flickr8k and Flickr30k datasets. Secondly, the preprocessing of the data must be done in a proper manner because as textual data goes, it contains punctuation, whitespaces, grammatical errors and so on. The link between visual importance and descriptions leads naturally to the problem of text summarization in natural language processing (NLP). In text summarization, the goal is to select or generate a summary for a document [4]. So using built-in Python libraries like NLTK can help overcome

S. P. Manay · S. A. Yaligar · Y. Thathva Sri Sai Reddy (✉) · N. J. Saunshimath
Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India

 
N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_43

this obstacle. Third and the final challenge is of detecting the objects in an image and organizing the sentence. If the data is not trained on properly, the output sentences will have recurring words or outputs that make no sense. Detecting the objects in the image is a bigger task because there are a variety of objects that can be in an image and one cannot develop his/her own model to detect these objects. So pretrained models like VGG16, Inception, ResNet50 and so on can help in this problem as it is trained on the famous ImageNet database. The applications of Image Captioning could include information retrieval (IR), guiding the visually impaired or even in producing robot human interactions. Here we present an application of Image Captioning by giving an environmental sense to the users through an Android Device. This is mainly targeted to the Visually Impaired to get the sense of what is around them. Most of the apps that are available for them are for navigation purposes or object detection. Our idea is to give them a sense of what is around them for them to feel it rather than navigate through.

### *1.1 Motivation for the Problem Statement*

The problem we would want to address here is the difficulty of the visually impaired in their day-to-day life. They face various challenges such as trying to navigate around to different places or trying to read materials because not everything is available in Braille. It may be easier in a closed space, but out in the real world in the streets, it is a lot tougher. They also would not want to get help from everyone as they do have the sense of not being taken care of a lot or being treated specially that is, they would want to be treated as every other person everywhere else. So with these challenges, it is difficult for the visually impaired to go about their tasks or daily life. We would want to address this situation to an extent by developing an app that the visually impaired can use to get a sense of the environment. There are apps for navigation that vibrate at an obstacle or give an auditory response. For these, the users will always have to keep their phones up which can be a bit inconvenient. So our app can be used whenever the user would want to. All he needs to do is unlock the phone, open Google Assistant and open the app through the command. Then the user can touch anywhere on the screen and take a picture of the surroundings. Then an audio will be played that tells the user what is in his surroundings through Image Captioning.

## 2 Datasets

The datasets used for this survey is the Flickr8k dataset. Flickr30k has a total of 31,783 images. MS-COCO is more challenging, which has 123,287 images [5]. The Microsoft COCO dataset and the Flickr30k dataset were not used as the dataset size was huge and there wasn’t enough processing power or capacity to get an inference from these datasets. The COCO dataset is a dataset containing approximately 83,000 training images and 41,000 validation and testing images with one caption for every image. The Flickr8k dataset is a dataset that contains 8000 images with 6000 images split into training set, 1000 images each for validation and testing set. Each image had 5 captions and the train-test split was already available in the form of filenames. The Flickr30k dataset is an extension to the Flickr8k dataset that has 30,000 images. It too has 5 captions per image. No train-test splits were available so it was manually split into approximately 19,000 training images, 6000 images each for validation and testing.

## 3 Previous Work

In [6] it is said that using the Flickr8k and Flickr30k dataset, the BLEU metric was used that could make the results that were achieved close to comparison with the best Image Captioning models. The BLEU metric is a metric that can be used to evaluate or judge the performance of translation, mainly machine translation. The evaluation is a modified version of precision. It is done by assessing the quality of the predicted text that is translated from one type of natural language. The training time of the network vastly depends on the hardware. The results can be improved by using better hardware. The authors divided their model into 3 phases: Image Feature Extraction phase is done through the VGG16 pretrained model. This in turn results in a 4096 vector elements which is carried over on to the Long Term Short Memory (LSTM) Layer. The Sequence processor is then used for processing this text input using a layer called the Embedding Layer. This too is connected to the LSTM Layer. The final phase is a Decoder phase which combines the above two phases and gives an output as text. The model was trained on a GTX 1050 GPU and used Keras for implementing the neural network.

In [7], the authors describe various methods for Image Captioning. A Retrieval-Based approach retrieves a sentence from a predefined pool of sentences. It is done through mapping the image to a Markov Random Field. Other methods include ranking through cosine similarities. In Template-Based method, the process of generating a description or a caption for any image, starts with identifying a particular collection of visual details. Then, these identified visual details are associated through sentence layouts or explicit grammar rules which can be combined to form a sentence. The final method is through Neural Networks such as multimodal learning, i.e., using a Convolutional Model and Language Model. Encoder-Decoder Architecture can be used as another method. An encoder neural network is where the neural network first tries to translate an image into a form of intermediate representation which then a recurrent decoder which is also a neural network takes this intermediate representation as its input and produces a word-by-word sentence. An extra layer called Attention Layer too can be added to improve the captioning which gives more importance to the salient features. While comparing all these Neural Network models using the BLEU metric, the Attention Layer method generated the best score among the other methods in both the COCO as just as for the Flickr8k dataset.

In [8] it is described on how to use ontologies that can be used to support annotation and searches in image collections. They have used four different ontologies like AAT, WordNet, ULAN, Iconclass. The subsequent RDF Schema documents are added into a tool with the assistance a parser called the SWI-Prolog RDF parser. This tool is then used to generate a user interface for annotation and also contains a search that is totally based on the specification of the RDF Schema. This tool has various features such as loading images and its collections, creating annotations, storing annotations in a format of RDF file, and lastly two types of image search techniques.

In [9] the authors describe the different ways of Image Captioning. Traditional Image Captioning works by taking patterns or histogram gradients which is sent to an SVM classifier. Other techniques are Deep Learning using CNN and RNN models. It is also mentioned about different methods of Deep Learning like Attention-based or Encoder-Decoder or Semantic Concept-based. The authors also introduce Reinforcement learning as a technique for Image Captioning. Using policy gradient techniques, a token is generated based on which the algorithm gets a reward for each state. GANs also can be used but it does not take into consideration the multiple captions of the training data. Using Dense Captioning which is used to obtain salient features of the image, the existing CNN approach can be improved. Semantic Captioning involves captioning based on semantic details. The results of the Image Captioning using datasets like Flickr30k and Flickr8k and COCO were inferred by using metrics like BLEU, ROUGE or METEOR. COCO dataset showed the best results due to their vast dataset and different examples.

# 4 Implementation

## *4.1 Image Captioning*

The model was implemented with a help of the VGG16 pretrained model. The VGG16 is a neural network developed by Oxford which contains 16 layers of neurons. It contains various layers like Convolution layer, MaxPooling layers, etc. This part of the network deals in extracting image features. This is required so that the objects present in the image are detected. Now to this pretrained network, another network is added to learn the captions for every image. As mentioned before each image comes with 5 captions, so the model learns the input which is a dictionary data structure containing the image features as key and preprocessed text as its value. By preprocessing the text, it is meant as removing punctuation and whitespaces, converting to lower case. In Fig. 1, the right part or area of the model is for the object recognition or extracting features of an image. The left part or area of the model is for the network to learn the captions. This was implemented using both LSTM and GRU. Since GRU showed better results in this particular model architecture, GRU is used as the model

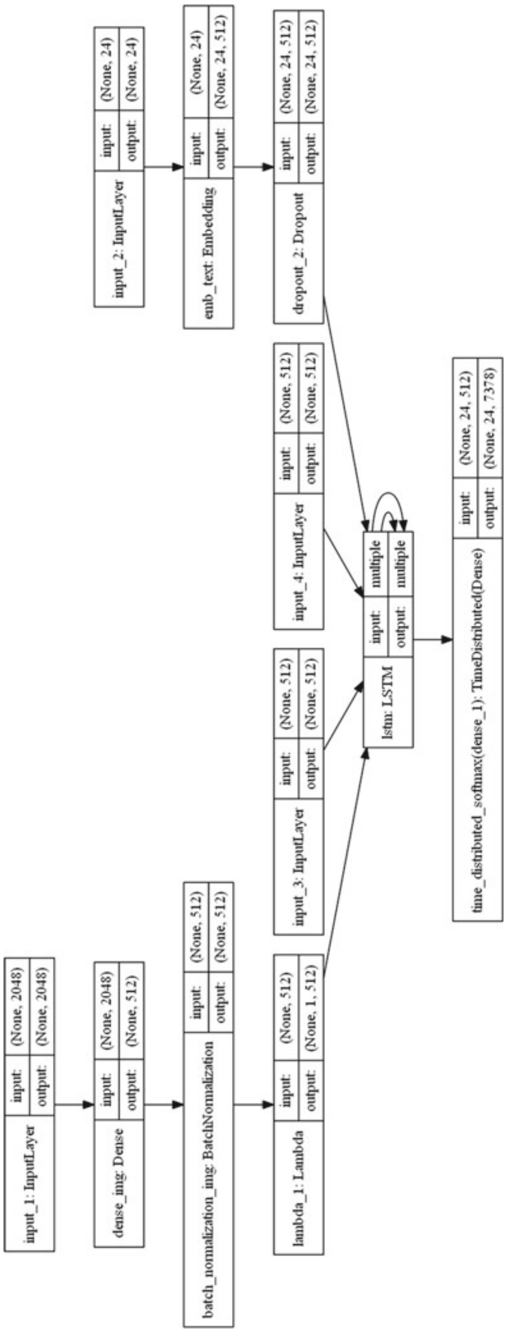



**Fig. 1** Model architecture

to display the results. Below is a description of each layer in the model architecture. Input Layer: The Input Layer is to accept the input for the neural network and pass it on to the other layers. There are two kinds of input layers, one is used for the images. The second kind is for the textual data. Embedding Layer: Since every neural network works on numerical data, the captions must be converted to a set of tokens or vectors with each token representing each unique word in the caption. Using Keras's Tokenizer library, the captions are converted to a number of tokens that is passed on to the network to the Embedding Layer. Now, the Embedding Layer learns these tokens so that it can predict the caption for a test image.

Dropout: This layer is to deactivate a certain percentage or a certain number of neurons in a layer to prevent the data from overfitting.

Dense Layer: The Dense Layer is where all neurons in the network are connected to each other in a 'Dense' manner.

(1) LSTM: LSTM is a solution used in Recurrent Neural Networks which stands for Long Short-Term Memory. LSTM is one of the solutions to solve the problem off the short-term memory problem in RNN. LSTM makes use of certain gates that allow information to either be passed through or be discarded. The cell state of the LSTM, carries important information or facts throughout the entire processing of a sequence. So the information or data that is from the earlier time when the information was processed can move to a later time, solving the problem of short-term memory. As the cell state keeps going on, information gets into and out of the cell state through neural networks called gates. The gates can be told as just different neural networks that help in deciding which or what kind of information should be allowed into the cell state. Based on these, the gates will be able learn which data is important to be kept or discard during the training process [10]. The components are:

Sigmoid: It is one of the activation function that standardizes the values in the range of 0 and 1. It is useful for the forget gate since anything multiplied by 0 is 0.

Forget Gate: The Forget Gate is used to decide which information is to be kept or discarded.

Input Gate: The Input Gate here is utilized to refresh the cell state. The data of the previous hidden state and the present data is given to the Sigmoid function which then standardizes the qualities between the scope of 0 and 1.

Cell State: For updating the Cell State, the following steps occurs: First the cell state goes through a mathematical operation which is pointwise multiplication. The forget vector and the cell state are pointwise multiplied. This creates an eventuality of dropping the values present currently within the cell state. Then, this output which is generated from the input gate, is taken and pointwise added. This will update the cell state to new values. Output Gate: The output gate here makes the decision on which information should be the passed on as the next hidden state (Fig. 2).

(2) GRU: GRU stands for Gated Recurrent Unit. It is part of a new generation of RNN which works similar to the LSTM. It only has two gates, i.e., reset and update gate. GRU is lightweight compared to LSTM. It evaluates faster than LSTM. But GRU is more suitable for shorter sentence vectors while LSTM has always been the standard method for any RNN tasks. GRU as well is one of the methods as a solution to the problem of short-term memory of the RNN. GRU was first established in 2014

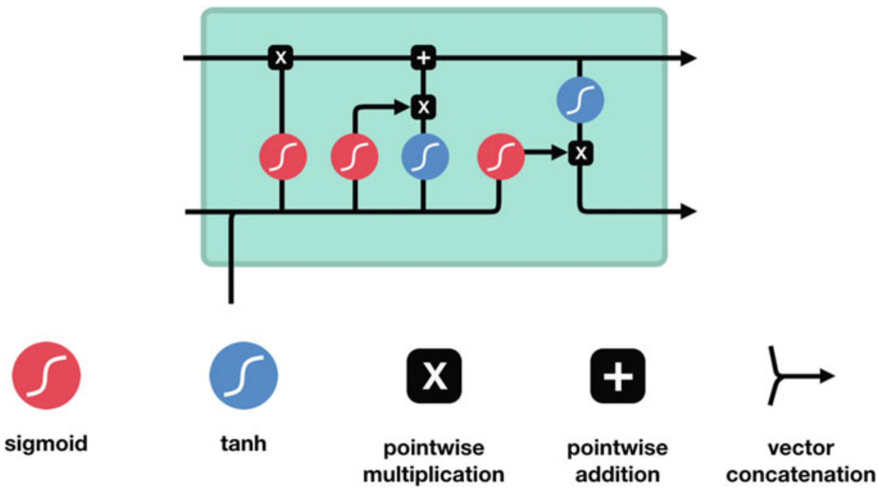



**Fig. 2** LSTM architecture

and some applications use GRU due to their speed and less computation intensive processing needs when it is compared to the LSTM.

Reset Gate: The Reset Gate is utilized to choose the amount of the past data is to required be forgetten.

Update Gate: It is almost the same as the Forget gate present in the LSTM method to decide which data it to be kept and what type of information is to discarded (Fig. 3).

### 4.2 *Working of Image Captioning*

The working of how the model learns to caption the images is simple. There are two features to be considered here. One of the features is from the image that is a image vector which consists of the objects and other information relevant to the image. The other feature is the text or captions that is relevant to the image. The textual features are tokenized into tokens of unique numbers using the Tokenizer library of Keras. Consider a data point as $Di = Xi, Yi$ where $X$ is the input for a data point $i$ and $Y$ is the target variable. So the input is an image and the caption is what is needed to be predicted. The prediction is done through the following steps: Initially as the first step, The data point of an image vector and the first word which is the input are used for the initial information to try and predict the second word. In the next input, the previous data that is the image and the two words are used to predict the third word. This continues until end of the caption. So at every step, a matrix of the image vector and the partial caption at each step is sent as an input to the LSTM cell. The word here is actually the tokenized tokens of each word. If the caption is unequal to what was the predefined sentence size, 0 s are padded. This way the model trains on every

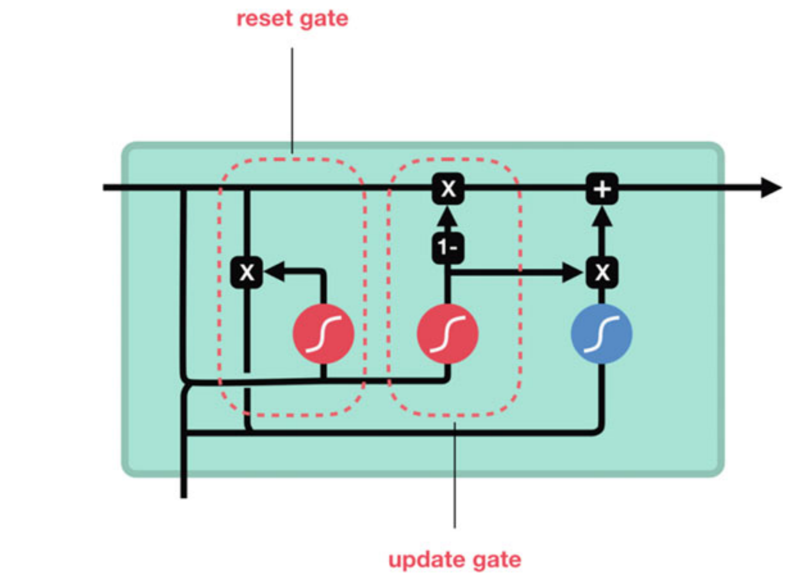



**Fig. 3** GRU architecture

image trying to predict the caption and learns by the error function and tries to give a better prediction in each epoch. This is the working of Image Captioning.

## *4.3 Android*

The Android application works as follows: The user touches the screen which directs the user to the camera where a picture can be taken. The user confirms the picture and is sent to the cloud which describes the image and an audio output is heard to the user. This is done through Asynchronous tasking of the tasks. The load on the Android application will be more due to the image depending on the Camera Megapixel density. Higher the camera density bigger will be the image size. So this could put a load on the user's application. To solve this, Asynchronous tasking splits the process into multiple threads so that the application does not work on a single thread and have a process overhead. This happens in the background of the application. After the image is taken the image is sent to the cloud. AWS EC2 instance is used as the cloud platform. A Flask REST API resides on the server that will be running on a GPU. Since Image Captioning is a highly compute intensive task, a GPU instance was used. The API keeps running and when the image is sent to the URL that is running on the server, the API processes the image and sends a response back to the client in JSON format. Using a library called OKHttp to handle the request and response, the value of the JSON response can be decoded. Post execution of this, the application is redirected to a screen that displays the image, has a textual description and also the audio that is being played.

## 5 Evaluation Metrics

For evaluation of how good the predicted captions were to the actual caption, accuracy would not be a good metric. This is because accuracy is where the predicted output exactly matches the actual output. In case of Image Captioning and textual outputs, though the predicted output may resemble the image, there may be a difference in the actual and the predicted caption. So to solve this, a metric called BLEU metrics is used. BLEU is an abbreviation that represents Bilingual Evaluation Understudy which is a calculation that contrasts the nature or quality of the text and a reference sentence. The BLEU score measures the modified n-gram precision of machine generated sentences with respect to human generated sentences [11]. The cornerstone of our metric is the familiar precision measure [12]. It evaluates by making a comparison of the candidate sentence which is the predicted sentence with a set of reference sentences, which are present in the dataset, in *n* grams. *N*-grams is a probabilistic model in which a collection of contiguous sequences of '*n*' items from any given sample or text is described. BLEU deals for the reference sentence by counting the corresponding *n*-grams from the candidate sentence. The comparison is made without considering word order. BLEU-1 uses unigram to compare with the reference sentence. BLEU-2 is a bigram comparison to be compared as a pair of words. BLEU-3 is a 3-g comparison with the reference sentence and the BLEU-4 is a 4-g comparison.

## 6 Results

This section shows the results of the trained Image Captioning model and the User Interface of the Android application (Fig. 4).

The model is able to correctly identify that there are two dogs in the picture and also describes its respective colors that is brown and white. It also is able to recognize the background that tells that the dogs are running in a field (Fig. 5).

The UI consists of a voice command that tells that the user can tap anywhere on the screen to take a picture and once the picture is taken and the output is heard. The UI is made with basic XML and Java language and works on any Android device.

**Fig. 4** Result 1


## 7 Conclusion

In this paper, the idea of using technology to help the visually impaired get the sense of their environment was discussed. The purpose of the application is achieved as the application focuses on enhancing the ability to get the sense or feel of the surroundings for the visually impaired. It also makes it efficient as it runs smoothly since the model is deployed on cloud reducing the workload of the Android application. The user can experience their surrounding at the comfort of their Android devices. The image captured is captioned which in turn is narrated via audio making it user-friendly. Several methodologies have been implemented in making the application a success. The Android interface is designed in such a way that the visually impaired person can easily use it. The efficiency and usability are counted as a priority as it is used by the visually impaired. Hence the application serves for the betterment of the visually impaired. Further, explanation of the datasets used and its specifications

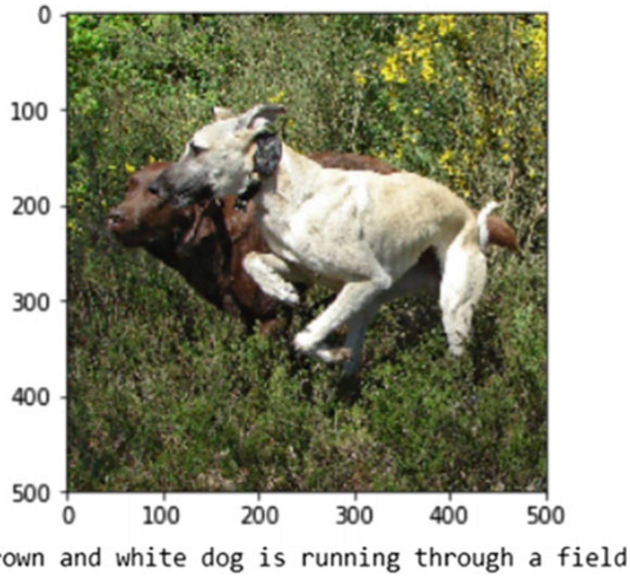



**Fig. 5** Result 2

were discussed. The approach was discussed as well. BLEU was the metric that was used to evaluate the results of the model. As part of the results, the images with the predicted captions were shown.

## References

1. L. Srinivasan, D. Sreekanthan, A.L. Amutha, Image captioning—a deep learning approach. Int. J. Appl. Eng. Res. **13**(9), 7239–7242. ISSN 0973-4562 (2018)
2. S. Bai, S. An, A survey on automatic image captioning. Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian District (Beijing, China, 2018)
3. L Hollink G Schreiber B Wielinga J Wielemaker 2003 Semantic Annotations of Image Collections University of Amsterdam Amsterdam, Netherlands
4. M.Z. Hossain, F. Sohel, M. F. Shiratuddin, H. Laga, A comprehensive survey of deep learning for image captioning. Murdoch University (Australia, 2018)
5. A. Karpathy, L. Fei-Fei, Deep visual-semantic alignments for generating image descriptions. CVPR (2015)
6. I. Sutskever, O. Vinyals, Q. Le, Sequence to sequence learning with neural networks. NIPS (2014)
7. Q. You, H. Jin, Z. Wang, C. Fang, J. Luo, Image captioning with semantic attention. CVPR (2016)
8. A. Farhadi, et al., Every picture tells a story: generating sentences from images, in *European Conference on Computer Vision* (2010)
9. G Kulkarni 2013 Babytalk: understanding and generating simple image descriptions IEEE Trans. Pattern Anal. Mach. Intell. 35 12 2891 2903

10. V. Ordonez, G. Kulkarni, T.L. Berg, Im2text: describing images using 1 million captioned photographs. Adv. Neural Inf. Process. Syst. 1143–1151 (2011)
11. K. Papineni, S. Roukos, T. Ward, W.-J. Zhu, Bleu: a method for automatic evaluation of machine translation, in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (2002), pp. 311–318
12. M. Phi, Illustrated guide to LSTM's and GRU's: a step by step explanation. Medium (Sept 2018). https://towardsdatascience.com/illustrated-guide-to-lstmsand-gru-s-a-step-by-step-explanation-44e9eb85bf21

# IoT-Based Water Quality Analysis and Purification System

**Ashutosh Singh, Akihil Ranjan, Nikhil, Manish Kumar Singh, Veda S. Nagaraja, and S. Raghunandan**

## 1 Introduction

Poor quality of water spreads illness in society. Poor quality does not only hampers the health of humans but also the life of fellow creatures of the earth. According to the Water Resource Information System of India, 2017, nearly 5 million people die due to waterborne diseases [1]. Agriculture plays a major role in depleting the quality of the water despite the industrialization and other urban-runoff since the fertilizers used by the farmers are washed away with the soil and gets deposited into the various water bodies. Modern dumped items from the urban cities are washed into the different water bodies nearby them thus providing significant damage to the quality of the water bodies. These polluting influences enter the natural way of life and accumulate until they reach harmful levels, in the long run hampering the life of the aquatic animals and significantly damaging the water quality. Data obtained from Central Ground Water Board, 2017 suggest that a rise in the temperature of water lowers the dissolved oxygen content and upsets the balance of aquatic life in water. Thus above-mentioned factors indicate the analysis of water quality and its subsequent purification makes it essential.

Real-time water quality monitoring and its subsequent purification based on the determined water quality parameters have always been a great challenge for different public health engineering departments and various households. Bad water quality spreads diseases, causes death and affects socioeconomic progress. Water Quality Analysis (WQA) data are not easily available to the general public in India. Advanced sensors-based technologies can be utilized in order to acquire information on water quality parameters and purify the given water accordingly. But these systems proved

A. Singh (✉) · A. Ranjan · Nikhil · M. K. Singh · V. S. Nagaraja · S. Raghunandan
Electronics and Communication Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka 560064, India



N. R. Shetty et al. (eds.), *Emerging Research in Computing, Information, Communication and Applications*, Lecture Notes in Electrical Engineering 789,
https://doi.org/10.1007/978-981-16-1338-8_44

to be very expensive for a common man in India. The proposed Water Quality Analysis and Purification (WQAP) system is of low cost, affordable and continuous real-time testing along with purification of water can be carried out at homes, colleges, hostels and any public establishments to ensure that clean and potable drinking water is available to the all irrespective of their socioeconomic status.

The point of this endeavour is three-layered. One is to give an itemized study of the latest work done in the territory of savvy water quality checking, it's application, innovations in this area, types of sensors utilized and so on. Second, is to introduce a minimal effort, less mind-boggling savvy water quality observing framework utilizing an Arduino board interfaced with a versatile application on screen to visualize the pH, Dissolved Oxygen, Conductivity, Turbidity, Temperature of the water. The system includes many graphical interfaces which facilitate to alert the user about the water quality parameters and keeps a regular check of the incoming water quality. The third venture of this project is the purification system which purifies the water based on a three-stage purification system centred on the water quality parameters provided by the measurement kit. The Purification system involves three phases, namely Primary phase, Secondary phase and Tertiary phase. The Primary phase involves the removal of all the heavy particles, grease, oil and dirt. Then in the second stage, it involves the removal of the Biochemical Oxygen Demand (BOD) by the application of a Dissolved Oxygen tank which continuously increases the oxygen concentration in the given water. The secondary system also consists of a High Rate Solid Contact Clarifiers Tank (HRSCC) tank and a filter membrane that chlorinates the water and passes through the membrane to removes all the small tainted particles. The tertiary segment contains Ultra Violet (U.V.), Reverse Osmosis (R.O.), Total Dissolved Solids (T.D.S) and multigrade filter, which makes the given water potable for drinking.

All three stages of purification undergo continuous monitoring with the measurement kit in order to incessantly check the parameters of the water and decide how much more selected filtration is required to meet the required quality of the water. The system is overall assimilated with an Internet of Things (IoT) by using a compatible Arduino board with a Bluetooth module to interface it with a mobile or computer to monitor the whole process seamlessly and get real-time results of water quality along with purified potable water.

## 2 Background Theory

From the underlying review and combination of water history and development at the Montreal meeting in 1997, it turned out to be evident that there is a need to reconceptualize mankind's history inside another hypothetical plan—one that would empower us to dig chronicled hotspots for a more profound comprehension of the powers and elements that have moulded and controlled human social orders in various environments [2]. Water quality's general job is a critical factor in accomplishing the Sustainable Development Goals (SDGs). Water quality is chiefly influenced by microbial and non-microbial boundaries. The greater part of the boundaries can be broken down rapidly by utilization of electronic sensors at a reasonable rate and in lesser time when compared to the monotonous natural and compound procedure. The boundaries incorporate broken up oxygen, pH, turbidity, temperature and all-out dissolved solids (disintegrated solids).

There are three notches of wastewater treatment: primary (first stage), secondary (second stage) and tertiary (or advanced).

## *2.1 Dissolved Oxygen*

**Dissolved oxygen (D.O.)** Amount of oxygen disintegrated (and thus accessible to support marine life) in a waterway, for example, a lake, pond, or stream. D.O. is the most substantial and promising parameter of the water body and has the ability to help an intelligent sea-going biological system of plants and creatures. The measure of disintegrated oxygen required fluctuates from animal to animal. Scavengers, crabs, shellfish and worms need negligible measures of oxygen (1–6 mg/L), while shallow water fish need more significant levels (4–15 mg/L) [2].

Dissolved Oxygen in a liquid is the free Oxygen ($O_2$) known as the free non-compound oxygen, which we can say exits in a divalent form in the water. The amount of dissolved oxygen in water determines the health of a water body. Dissolved Oxygen plays the most significant in support of life of aquatic living beings. A too low or too high content of water dissolved oxygen may harm marine creatures.

## *2.2 pH*

pH is defined as the capability of a solution to hold hydrogen ions. The A pH is a measure of hydrogen ion concentration, a proportion of the acidity or alkalinity of an answer. pH is basically the concentration of $H^+$ ions in a liquid. The pH is the most day to day term and a common man hears for if a discussion on water quality happens. pH is one of the fundamental parameters in determining the quality of water. pH is generally a logarithmic scale from 0 to 14. pH determines the acidity or alkalinity of a solution. pH plays a dominant role in deciding hard water qualities, as we all know that hard water is a very problematic situation in India.

## *2.3 Turbidity and Total Suspended Solids(TSS)*

Turbidity is the cloudiness or opaqueness observed in a liquid when seen through the projection of light. Turbidity is yet another essential water quality parameter that helps to determine the purity of water. The turbidity is the number of colloids present in the water. The more the number of colloids, the more is the turbidity of water. Turbidity is measured in terms of NTU. NTU is an S.I. Unit given by the international community of measurement of liquids, gases and solids. NTU is Nephelometric Turbidity Unit, which is the standard unit for measuring the amount of haziness in the fluid. NTU ranges from 0.1 to 25, 0.1 considered being free from all colloids whereas 25 considered to be as dirty opaque sewage water. Perfect pure water has a turbidity range within 1–2 NTU. In context, turbidity and hard and total suspended solids (TSS) are associated. Regardless, turbidity is a prompt estimation of the hard and fast suspended materials in water. Turbidity and all-out suspended solids suggest particles present in the water section. Outright suspended solids are a full-scale sum estimation of solid material per volume of water. This also infers TSS is a specific estimation of each and every suspended solid, characteristic and inorganic, by mass. TSS fuses are settle-able solids and are the quick estimation of the outright solids present in a water body [3].

### 2.4 Primary Treatment

The primary stage is the first stage in purification of water and also consists of a preliminary stage. Primary step removes nearly 60% of suspended particles (massive, floating like materials in water) using a screening chamber and 35% of BOD (Biochemical Oxygen Demand) using different methods. Primary treatment involves four prominent processes. The process is the screening, in which the sewage is sent into a bar screening chamber where a mesh is placed and the sewage water is passed through it. After the bar screening chamber, it goes to the grit chamber where all the bigger suspended particles are removed and the turbidity is checked. From the grit chamber, it enters into the oil skimmer chamber where all the greasy and oily contents in the sewage are removed using a skimmer. The collected sludge is simultaneously passed into the disposal tank and the first stage cleared water is sent into the second stage.

### 2.5 Secondary Treatment

Secondary treatment is also known to be an Auxiliary treatment process and involves the removal of microbial organisms by increasing the dissolved oxygen content in the water with the help of an aeration chamber. In the secondary treatment process, the water purified from the first stage is collected into a chamber and then sent into an aeration tank where the water is rotated at high speed in order to increase the content of dissolved oxygen. The dissolved oxygen content is increased to a certain level and then water with a high amount of D.O. is passed through a filtration membrane in order to remove all the fine particles of the water, after passing from the filtration membrane it is sent to chlorination tank where the water is chlorinated and all the infectious contents are removed.

### *2.6 Tertiary Treatment*

The main objective of the tertiary stage treatment of water is to remove the microbial and non-microbial hazardous contents of the water and make it suitable for drinking and household purposes. The tertiary treatment removes the nitrates, phosphates and soapy parameters of the water and also ensuring that the water is soft after the treatment.

Treatment choices in tertiary treatment rely on the attributes of gushing after auxiliary treatment and what sort of water is required at the end of the process. For instance, on or off chance that one needs consumable water, at that point filtration and sterilization are to be done on the wastewater.

## 3 System Architecture

The projected system has two main parts, the water quality measurement kit and a three-stage purification system installed with an arduino connected with bluetooth in order to connect with the internet to send and receive the commands at any location.

The Measurement Kit consists of many different sensors, namely, Dissolved Oxygen, Temperature, pH, Conductivity and Turbidity. Since every electrode is associated with an interface and this interface with the Arduino gathers simple information and changes over it into the electronic structure. The data is then processed by the Arduino and is sent to the P.C. to display the results of the quality parameters observed in water. The various analogue data collected by the electrodes of different sensors are converted into suitable S.I. Form of representation by applying the algorithm developed for each and every sensor. The user then reads the data of the water quality and sends the command to the purification system on how much purification is required.

The purification system consists of three stages which include bar screening chamber, aeration tank, sludge removal chamber and then Ultra Violet (UV), (Total Dissolved Solids) TDS and Reverse Osmosis (RO) control rods for making the water suitable for drinking or household purposes.

Figure 1 gives a brief introduction of how the integrated system of water quality measurement kit and purification system works. The Measurement kit is connected to a mobile application through the cloud, which sends all its processed values from the different sensors to the developed cellphone application. From cellphone the user can interpret the quality of the water and decide what to do next, i.e. how much or at which type of purified water is required, whether for household purposes or for drinking

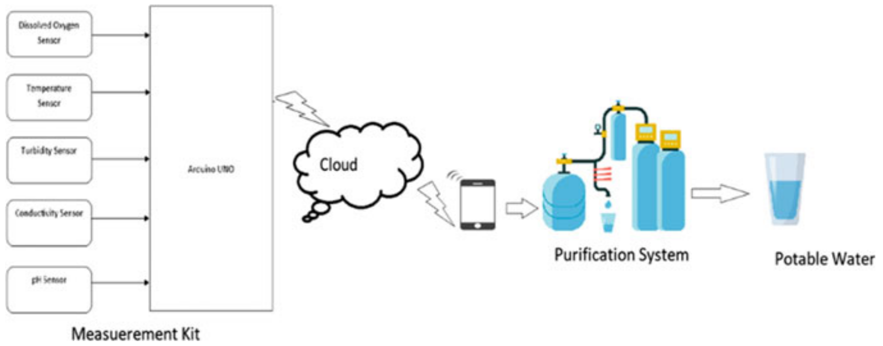



**Fig. 1** Integrated system block diagram

and based on his/her requirement can send the command to the purification system to purify the water accordingly based on the three-stage process of the purification.

### 3.1 Measurement Kit

Water quality monitoring is viewed as the collection of data regarding the quality parameters of water at a predefined location and at customary spans so as to give information that might be utilized to give the correct standards of water and how much purification does it require. The measurement kit comprising of different sensors, namely Dissolved Oxygen, Temperature, pH, Turbidity is interfaced with the Arduino, which in turn sends the collected and processed data to the cell phone via Bluetooth module.

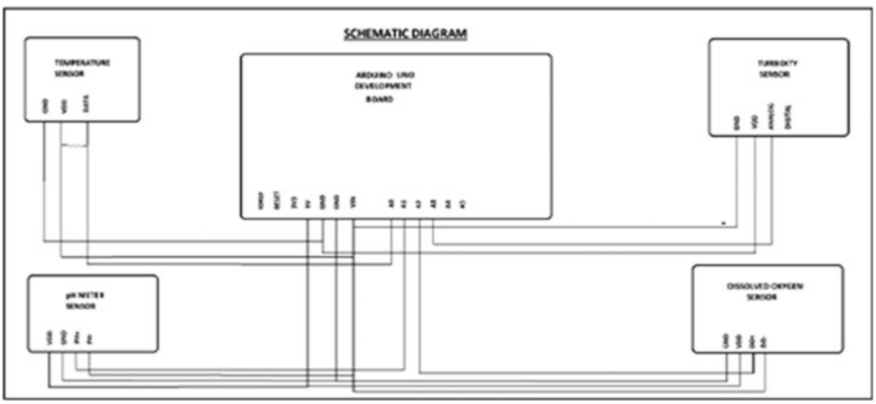
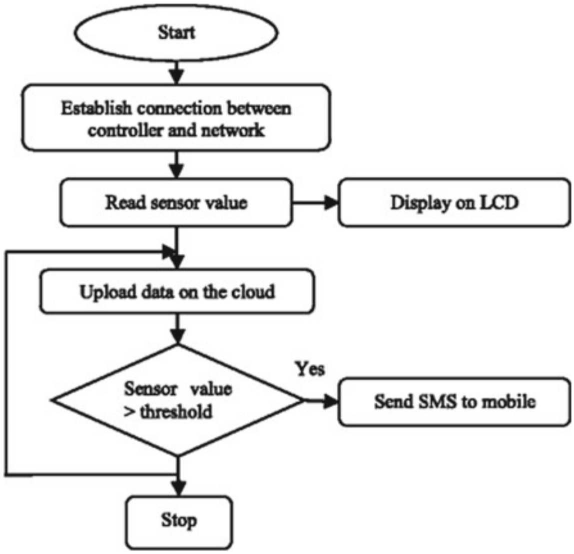



**Fig. 2** Schematic Diagram of measurement kit

**Fig. 3** Flow chart for sensor data storage




Circuit connection diagram of the measurement kit is shown in Fig. 2. The figure gives the required connections in order to interface different sensors with the arduino. All the sensors are legitimately interfaced with the controller since the proposed architecture is to screen household water quality.

The information gathered from the sensors from the desired solution is collected and compiled by the controller and is then sent to the cloud and from the cloud, the data is obtained onto the cellphone's application. Respective commands can be given by the cell phone for activating the specific stages of purification based on the quality parameters of water obtained from the controller.

Figure 3 presents the flowchart of steps involved in the process of data storage.

## *3.2 Purification System*

The primary focus of the water cleaning process is to bring out water that is fitting for explicit purposes, for example, drinking, clinical, pharmaceutical, synthetic and mechanical purposes. Filtration process diminishes the centralization of the suspended particles, microorganisms, green growth, infections and parasites alongside the scope of the disintegrated and particulate issue.

Figure 4 shows the detailed process of primary and preliminary treatment of water. The wastewater is collected into a receiving container and is sent to the bar screen chamber for removal of large suspended particles of dirt, etc. From it, the sample water is passed to grit chamber for removal of medium-sized particles and from there it is sent to oil and grease chamber for removal of oil and greasy substances from

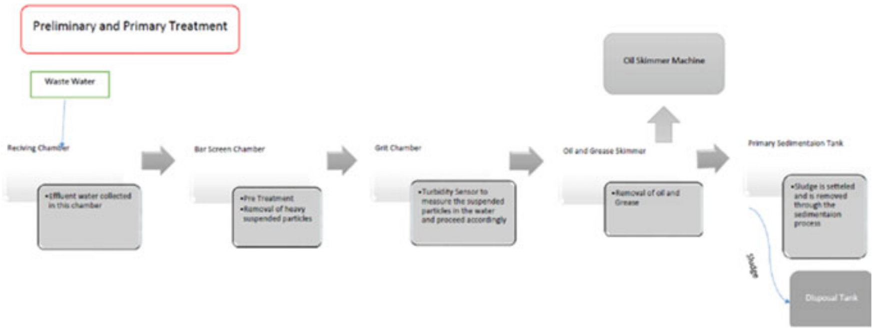



**Fig. 4** Flow chart—primary treatment

the water with the help of an oil skimmer. The Primary stage purification comes to an end at this point and the activated sludge is removed through a disposal tank.

Auxiliary treatment of wastewater chips away at a more profound level than essential and is intended to significantly corrupt the organic substance of the loss through oxygen consuming natural procedures.

Figure 5 depicts the working of secondary treatment. This stage involves aeration tank, which helps in levelling up the dissolved oxygen content of water by reducing the bio-organic waste. The aeration tank is followed by a PAC addition chamber which adds powdered activated carbons and rotates the water at high speed, consequently removing all the harmful BOD components. Water is then sent to the HRSCC tank followed by a filtration membrane. Simultaneously a water quality measurement kit is also placed in order to check quality parameters.

The tertiary treatment stage inside the wastewater board framework readies the wastewater for definite use and how altogether it should be dealt with relies upon the wellspring of the wastewater and what it will be utilized for. Figure 6 shows the stream graph procedure for tertiary treatment. Treatment alternatives in tertiary treatment rely on the qualities emanating after optional treatment and what sort of

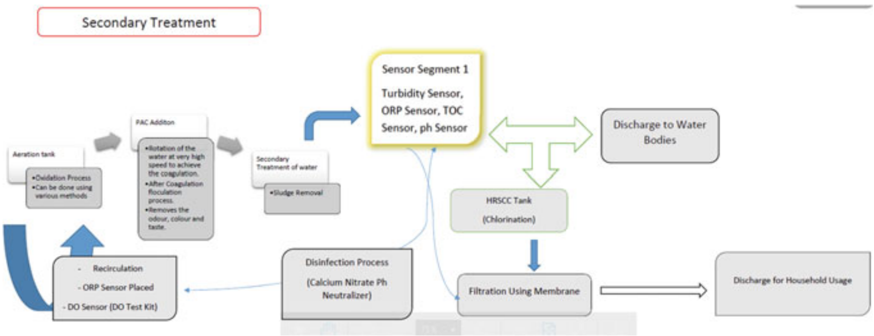



**Fig. 5** Flow chart—secondary treatment

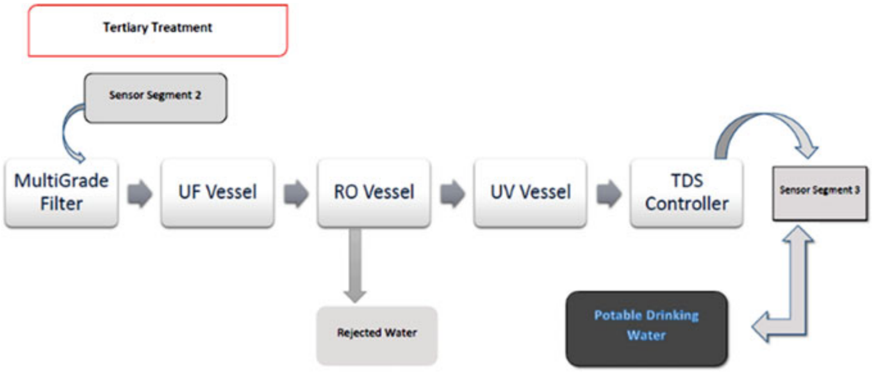



**Fig. 6** Flow chart—tertiary treatment

water is required towards the finish of the treatment. For instance, in the event that one needs consumable water, at that point filtration and sterilization are to be done on the wastewater. As it's obvious from the chart the water got from the auxiliary stage is sent to a multistage channel from that point it's gone through U.V. to pulverize microorganisms and debase natural contamination. Ozone gas ($O_3$) is framed by going oxygen through a high voltage electric field. From UV it is sent to RO and afterwards TDS for keeping up ideal supplements in the water to make it potable.

## 4 Implementation

The proposed system is implemented using Arduino Uno interfaced with all the different sensors and then integrated with the purification system with the help of solenoid valves for controlling the flow of purification process.

A temperature sensor is a thermocouple (T/C) or RTD device that calculates an approximation of temperature through an electrical sign. This device uses two metallic terminals to voltage difference with a directly change in temperature. The DS18B20 sensor, which is stainless steel is a wired and waterproof kind. Its unique one-wire interface has an easy to communicate frame with other devices. It changes temperature over the digitized word of 12 pieces in a maximum of 750 ms. As it draws power from the information line hence doesn't need any outside force. [4].

pH electrode is a straightforward and quick gadget to gauge the alkalinity and acidic nature of any liquid. A pH meter can be seen as a voltmeter that shows the electrical difference amount between a pH terminal and KCl reference terminal and produces the output of the difference as its pH of the water in which the probe is submerged.

pH meter gives an incentive concerning how alkaline or acidic it is. pH sensor will measure the group of charged $H^+$ ions particles. Acids become stronger if all the charged $H^+$ ions are grouped together emphatically. On-trade balance is set up between the fixed destructive destinations on the glass surface of pH electrode and $H^+$ , with an expanding number of accuse sets of $H^+$ happening as its movement in the reaching arrangement increments. This equilibrium can be communicated by

$$H^+/-O-Si- \rightarrow H^{++}-O-Si- \quad (1)$$

The arrangement having pH esteem one will be profoundly acidic and with pH esteem 14 will be exceptionally fundamental. The sharpness and alkalinity of any arrangement rely on the concentration of hydrogen particles ($H^+$) and hydroxyl particles ($OH^-$) correspondingly. An impartial arrangement as distilled water has pH 7.

$$Ag_{(S)}\left|\ AgCl_{(S)}\right|Cl^-_{(aq)}\left|\right|\ H^+_{(aq,\ unknown)}\left|\right|H^+_{(aq,\ fixed)},\ Cl^-_{(aq)}\left|\ AgCl_{(S)}\right|Ag_{(S)} \quad (2)$$

The mix terminal allows the estimation of both the interior and the outside layer surface prospects, which, as we saw above, is related to the course of action pH by the Nernst condition. The speculative potential over the glass layer changes by 59.2 mV for each unit change in game plan pH. A blend terminal cell can be addressed by the shorthand documentation underneath [5].

Figure 7 shows the configuration diagram of the pH and temperature sensor with the controller.

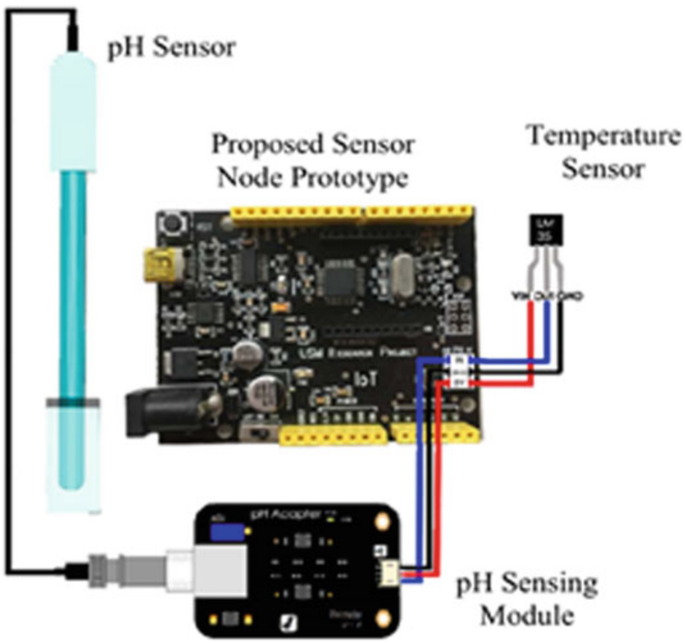



**Fig. 7** Sensor configuration of pH and temperature sensor

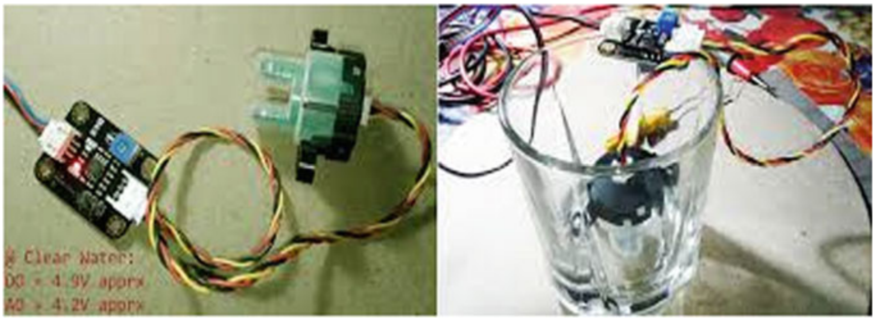



**Fig. 8** Turbidity sensor interface configuration

Turbidity is accumulated in water due to the following reasons:

- Gravel, sand, mud and silt
- Bacterial germs
- chemical precipitation.

It is necessary to treat water and check quality parameter of household water use such that it always experience some problem. For example, High turbidity water blocks the canals and channels in the stormy, rainy season and prevent them from working efficiently.

This Arduino based sensor checks different water quality parameters by calculating the degree of turbidity in water. It utilizes normal light to separate the suspended particles present in water by computing dispersing and transmittance rate, which changes with change altogether suspended particles (TSS). With TSS increases, turbidity increments with time.

Turbidity sensor has an operating Temperature of range 5°C to 90°C and voltage and current operating at 5 V and 40 mA maximum respectively with a supply of resistance (insulation) of 100 M at least and response time of less than 500 ms.

Figure 8 shows the image of the turbidity sensor interface configuration while developing the kit.

There are two key procedures for estimating Dissolved oxygen—galvanic and polarographic. Both probes utilize a terminal framework where the broken down oxygen response with the cathode to deliver current.

The Dissolved Oxygen is the most vital factor in the determination of water quality and concentration of water varies with the temperature of the water as well as its surrounding environment. Thus this also paves the way for measuring the temperature when D.O. is measured simultaneously in order to conclude the quality of water.

The centralization of Dissolved Oxygen (as estimated by its incomplete weight) is contrarily relative to iridescence lifetime, as appeared by the Stern–Volmer condition [6].

$$T_{\mathrm{O}}/T = 1 + k_{\mathrm{q}} * t_0 * \mathrm{O}_2 \tag{3}$$

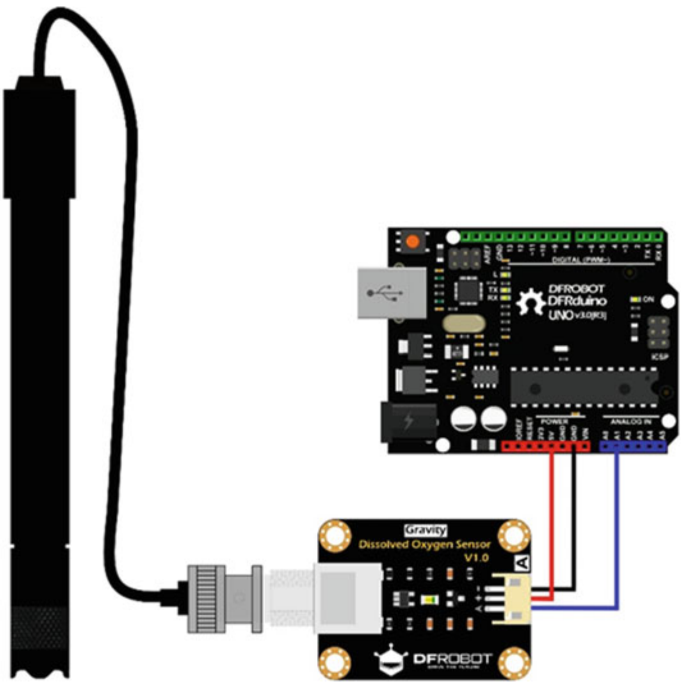



**Fig. 9** D.O. Sensor configuration [7]

$T_O$ Life of dye luminescence without oxygen
$T$ Life of luminescence with oxygen
$k_q$ Coefficient of Quencher
$t_0$ Luminescence life of the dye
$O_2$ Oxygen concentration as a partial pressure

Figure 9 depicts the sensor configuration of a dissolved oxygen electrode. The dissolved oxygen is based on dual point calibration. Temperature, pressure and salinity do not have any effect on the calibration (Fig. 10).

Dual point calibration is performed in order to achieve accurate readings below 1 mg/L [8].

The steps involved in dual point calibration are.

- Cautiously pull off and dispose of the top from the D.O. test.
- Allow the test to sit, presented to air until the readings balance out.
- When the readings have been balanced out issue the alignment order calculation in the sequential screen.
- After alignment is finished, you should see readings between 9.09 and 9.1 $\times$ mg/L, to see if temperature, saltiness and weight pay are at the default esteems.

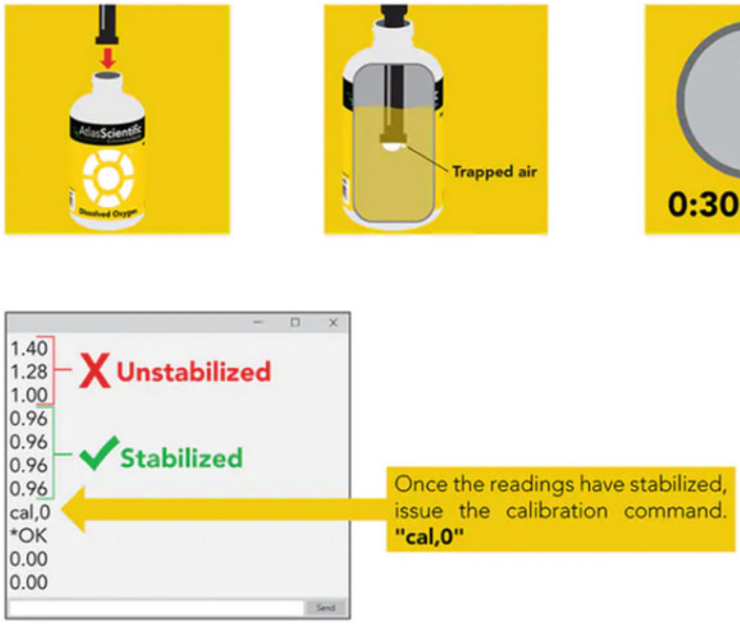



**Fig. 10** Dual point calibration

- After adjusting the D.O. circuit utilizing the "cal" order; place the test into the adjustment arrangement. Mix the test around to evacuate caught air (which could make readings go high).
- Let the test sit in the alignment arrangement until the readings settle. Note: Small development starting with one perusing then onto the next is standard.
- Once the readings have been stabilized, keep the calibration command "cal" to 0 in the serial monitor.

Figure 11 portrays the whole measurement kit system interfaced with all the sensors within a contained box. The red lights of the kit suggest that the system is on and ready to measure the water quality values and give the desired result.

The Purification System consists of a primary chamber, aeration chamber and PAC chamber and multigrade filter.

## *4.1 Primary Chamber*

The primary chamber is inbuilt with 20-micron filter cartridge incorporated with 12 V D.C. motor which takes the sewage from the water and filters all the heavy suspended particles from the water and sends the water into the aeration tank. Figure 12 shows the primary screening chamber consisting of a mesh to filter out large suspended

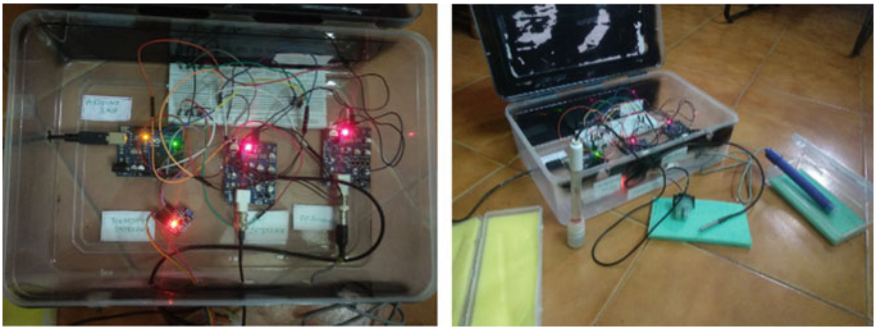


**Fig. 11** Developed measurement kit

**Fig. 12** Primary screening chamber

particles, also the chamber is inbuilt with the micron filter to remove the small tiny dirt and particles.

## 4.2 Aeration Chamber

Aeration chamber is used as the first treatment before any treatment this process annihilating Hydrogen Sulfide and Iron (spoiled egg smell) from the input. Air acts as an oxidizing agent for both Hydrogen Sulfide and Iron. It instantly changes from unfiltered Ferrous Iron to filtered Ferric Iron and also reduces Hydrogen Sulfide to Sulfur atom, which is then carefully flushed out from the water by a passage [7]. Figure 13 depicts the inside structure of an air circulation tank. The vacuum apparatus injects air into the tank and a small passage of compact air shapes in around 1/third piece of the head of the tank. As water enters the tank through the funnel at left, it hits an astound (of the three channels joined to the vent head, it's the short channel on the left) and washes out through the passage of compacted air. The water also

**Fig. 13** Aeration chamber

circulates air inside the tank before flushing out through the cylinder (called a riser) which receives into the bottom of the tank and flush out through the channel at the correct side and where mid-length tube in the centre is called Vent tube. It controls the profundity of the air pocket.

Figure 13 shows the developed aeration chamber consisting of an air pump incorporated with a 12 V DC motor.

### 4.3 PAC Chamber and Multigrade Filter

Activated carbon filters come into picture when expelling natural wastages and also while expelling chlorine (free) from water, such that after this it will be for appropriate use in the next phase. Taking the organics like humic and fulvic out from drinkable water and keeping chlorine in the water and reacting with the acids and trihalomethanes (Cancer-Causing agents) [9].

All actuated carbon channels expel an essentially higher measure of air toxins than customary carbon channels; anyway, there are a few qualifications. A few channels contain more actuated carbon than others and this can influence the limit with respect to pollution evacuation. The more enacted carbon is in a channel, the more contaminants it will get and the quicker it will adsorb. Figure 14 shows the PAC cartilage, which is incorporated in the purification system.

Figure 15 depicts the connection diagram of the tertiary segment consisting of sediment filter, PAC cartilage and U.V. sterilizer. The Solenoid valve is placed in the segment and the following connection has been designed in order to get connected with the other segments and then with the controller unit.

**Fig. 14** PAC cartilage

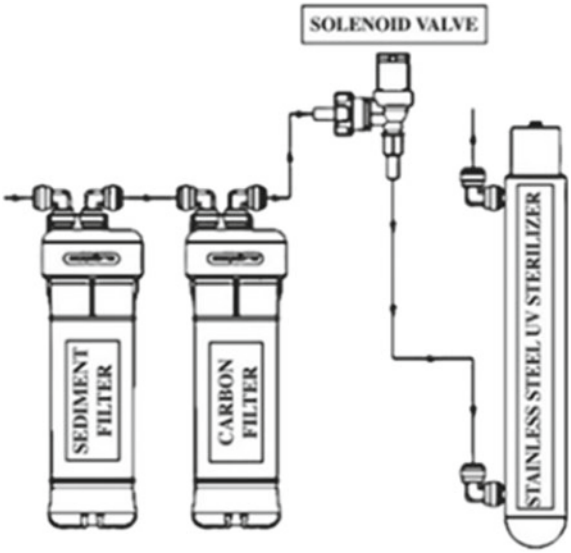



**Fig. 15** Connection diagram of sediment and carbon filter

# 5 Performance Evaluation

In our proposed water quality analysis and purification system, the water quality parameters were determined for various water samples taken into account and then sent for purification. Also, a chemical-based water quality check was performed in order to evaluate the performance of the water quality monitoring system as well as the purification kit. The real-time pc and mobile application results are shown in Fig. 16a.

Figure 16b depicts the personal computer-generated real-time result of the water quality parameters of the given sample of water.

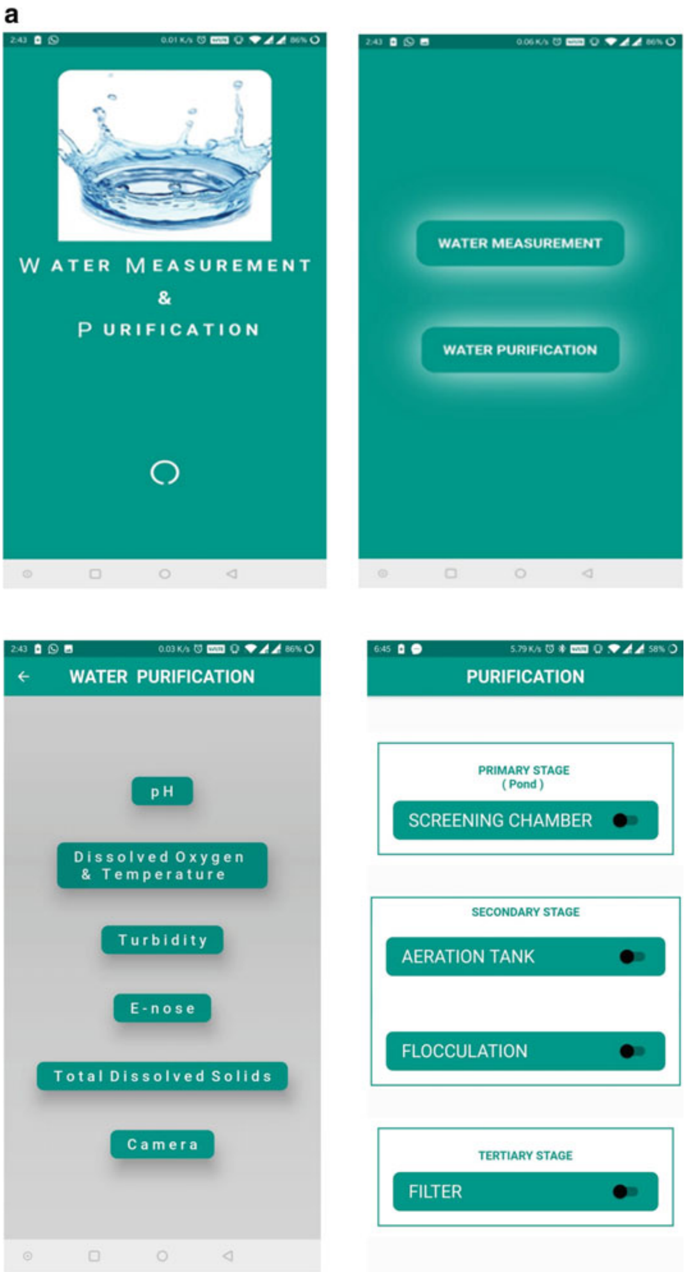



**Fig. 16** **a** Android application pages. **b** Real-time pc generated a result

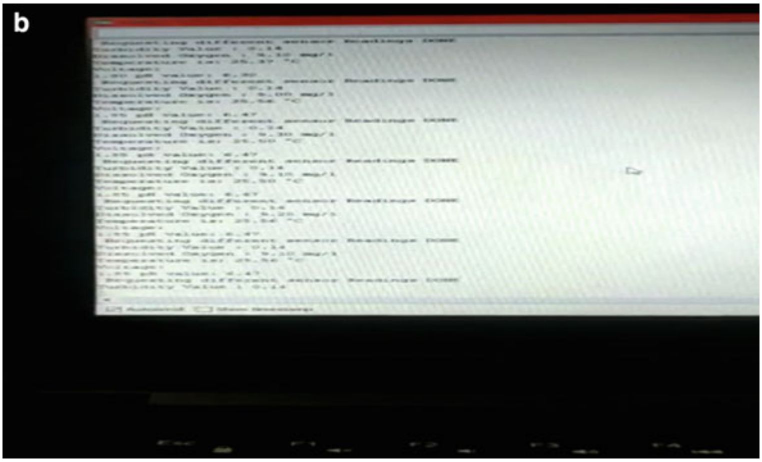



**Fig. 16** (continued)

The Measurement kit was validated at the Environment Laboratories, by performing a comparison of water quality parameters carried out with Chemical Process. The environment lab operators applied a chemical process for measuring the dissolved oxygen, temperature, turbidity and pH. Alongside our team performed the test for water quality parameters using the measurement kit. The sample considered for measuring the parameters was normal tap water.

Table 1 shows the result of the Chemical approach performed by the environment lab operators and Measurement kit results for the given sample of tap water (Table 2).

The above table shows the chemical report of bore well water. From this table, one can conclude that the values of the chemical report and the values obtained from the experimental analysis are almost matching with each other.

**Table 1** Tap water

| Parameters | Chemical approach report | Measurement kit report |
|---|---|---|
| Turbidity (NTU) | 3 | 3.34 |
| pH | 6.2 | 6.19 |
| Dissolved oxygen (mg/litre) | 7–8 | 7.4 |
| Temperature (°C) | 27 | 27 |

**Table 2** Bore well water

| Parameters | Chemical approach report | Measurement kit report |
|---|---|---|
| Turbidity (NTU) | 4.2 | 4.34 |
| pH | 8.2 | 8.23 |
| Dissolved Oxygen (mg/L) | 9 | 8.8 |
| Temperature (°C) | 24 | 24 |

## 6 Conclusion

In this work, an automated water purification cum measurement system was developed using a low powered Arduino board. The system is useful and time & cost-optimized simultaneously enabling us to monitor the water quality and have purified drinking water. The measurement analysis kit has been designed and successfully tested by measuring different water samples, be it sewage water or normal household or domestic purpose used water. The paper reported a handy IoT—based water quality monitoring and purification system, which could be easily installed in any home or any public establishment. The accuracy of the measurement kit suggests that it can be employed for the basic use of human beings and provide real-time parameters of water required at any instant of time with the least wastage of water.

The investigated system can monitor the quality of water automatically without the intervention of a human and it is considered to be low in cost. This system is used to keep the water from being pure by measuring pH, TDS, conductivity, dissolved oxygen, turbidity and temperature using wireless sensor network nodes and monitoring the status of purification system by internet using secured Web at home, offices, etc. According to that, the water quality monitoring becomes more economical, convenient and fast. This work suggests that water quality analysis and purification systems can be easily applied to homes, offices and schools and any other place. By building the described system in a smart building, the users will be capable of reading and analyzing the water usage specifications of the residents and thereby save a lot of water from wastage. It is also seen that the accuracy of measurement of the kit developed is more than 93% when compared with the results obtained through standard procedure.

**Acknowledgements** The authors would like to thank the Center for Nanomaterials and MEMS of Nitte Meenakshi Institute of Technology for sponsoring the research.

We wish to convey our special gratitude towards Dr Ramesh B R, Head of Department of Civil Engineering NMIT, Prof. Dr Vidyavathi N., Department of Civil Engineering, NMIT and Dr Erwin Reichel, Researcher, J.K. University, Austria for their valuable guidance for the conduction of this research.

We would also like to thank the students of Civil Engineering Department for their valuable support and help.

## References


1. Water history for our times, Person as author: Hassan Fekri, Document Code: SC2011/WS/5 , Collation:122 p. Language: English, Year of Publication:2011 (Online). Available: https://unesdoc.unesco.org/ark:/48223/pf0000210879.locale=en
2. Working group report (Brussels, 20–23 March 2014). Health effects of the removal of substances occurring naturally in drinking water, with special reference to demineralized and desalinated water. EURO Reports and Studies 16. Copenhagen: World Health Organization [WHO] (2014)
3. Editor-in-Chief A.H. Hassani, Director-in-Chief Majid Abbaspour (2020) Int J Energy Water Resour (IJEWR) 4(1)
4. DFRobot, "PH Meter(SKU: SEN0161)—robot wiki," 2016 (Onine). Available: https://wiki.dfrobot.com/PH_meter_SKU__SEN0161_
5. Waste Water Treatment Bodog Marinela, Derecichei Laura University of Oradea, Faculty of Environmental Protection, 26. Gen. Magheru St., 410048, Oradea
6. Fondriest Environmental Learning Centre (Online). Available: https://www.fondriest.com/environmental-measurements/measurements/measuring-water-quality/dissolved-oxygen-sensors-and-methods
7. DFRobot, "DO Meter (SKU: SEN0161)—robot wiki," 2016 (Online). Available: https://www.dfrobot.com/product-1628.html
8. Handbook for the operation of Water Treatment Works, Editor: Frik Schutte, Water Utilization Division, Department of Chemical Engineering, University of Pretoria TT 265/06 MARCH 2006
9. Environment Lab Manual, Environments Laboratory Civil Department, Nitte Meenakshi Institute of Technology (2019)
10. Arduino Project Hub Dissolved Oxygen Sensor Calibration 2020 (Online) Available:https://create.arduino.cc/projecthub/atlasscientific/arduino-dissolved-oxygen-sensor-calibration-69a4e2