Cognitive Science and Technology

Amit Kumar
Gheorghita Ghinea
Suresh Merugu *Editors*

Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing

ICCIC 2022, 27–28 December, Hyderabad, India; Volume 1

Springer

Cognitive Science and Technology

Series Editors

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

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

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

Indexed by SCOPUS

Amit Kumar · Gheorghita Ghinea · Suresh Merugu
Editors

Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing

ICCIC 2022, 27–28 December, Hyderabad, India; Volume 1

Springer

Editors
Amit Kumar
BioAxis DNA Research Centre Private Limited
Hyderabad, Andhra Pradesh, India

Gheorghita Ghinea
Department of Computer Science
Brunel University
Uxbridge, UK

Suresh Merugu
CMR College of Engineering and Technology
Hyderabad, India

ISSN 2195-3988 ISSN 2195-3996 (electronic)
Cognitive Science and Technology
ISBN 978-981-99-2741-8 ISBN 978-981-99-2742-5 (eBook)
https://doi.org/10.1007/978-981-99-2742-5

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

Making Cell-Free Massive MIMO Using MRC Technique

Sk. Khudsia Tabassum, D. Vishnuvardhan, G. Mamatha, and Fahimuddin Shaik

Abstract While demonstrating that partially or completely centralising signal processing at the central processing unit (CPU) improves spectral efficiency, this study advocates for the utilisation of maximum ratio trying to combine at each access point (AP) (SE). Cell-free massive MIMO is indeed a feasible solution for meeting rising user demands and great hopes in post-5G systems. Basic idea is to enable disseminated information access points (APs) to interconnect with all consumers on the system, probably via a single monolithic signal dispensation system. The purpose of the presented paper is for introducing the first thorough assessment of technology at various levels of AP collaborative efforts. The uplink efficiency of four distinct cell-free adoption and implementation is systematically explored using spectral fading and unrestrained linear processing. It did turn out that only by using MRC can it easily exceed both small cell as well as conventional mobile massive MIMO networks. As a result, it is the preferred method for operating CFM-MIMO systems. Nonlinear interpreting research also demonstrates that it makes only a minor difference.

Keywords MIMO · Cellular · Spectral · Wireless · 5G

Sk. Khudsia Tabassum
DECS, ECE Department, JNTUA College of Engineering (Autonomous), Ananthapuramu, Andhra Pradesh, India

D. Vishnuvardhan
Department of Electronics and Communication Engineering, JNTUA, Anantapuramu, Andhra Pradesh, India
e-mail: vishnu.ece@jntua.ac.in

G. Mamatha
Department of ECE, JNTUA College of Engineering, Ananthapuramu, India
e-mail: gmamatha.ece@jntua.ac.in

F. Shaik (✉)
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, Kadapa,, Andhra Pradesh, India
e-mail: fahimuddin.shaik.in@ieee.org

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_1

1 Introduction

The cellular network structure depicted in Fig. 1 is the traditional method for providing wide-area wireless communication services [1]. Each base station (BS) supports a distinct collection of user equipment (UEs). This network topology has been used for many years, because of relatively small size and cell and the application of more intelligent techniques for signal processing for interference reduction, spectral efficiency (SE) had also steadily increased [2]. Massive multiple-input multiple-output (mMIMO) has emerged as the principal 5G physical-layer technology. It is possible to boost the SE over conventional cellular networks through at least tenfold by helping to improve the BS hardware rather than establishing new BS stations. The SE gain is derived from a small array of 100 or more antennas on each BS, that is utilised for digital modulation schemes as well as, more accurately, to spatial and temporal multiplex different users equipment (UEs) employing the same time–frequency resource [3]. MIMO varies from conventional MU-MIMO in that every BS contains array elements compared to UEs within cell. Each BS can use interpersonal style processing methods, such as MRC inside the uplink, to reduce intervention between the same cell and other cells without mandating any BS cooperation [4]. As illustrated in Fig. 1, the mMIMO process also allows for positioning with spatially distributed arrays in each cell. This setup is very equivalent to the synchronised multi-point (COMP) as well as distributed antenna system (DAS) configurations with stationary, disconnected cooperation clusters [5]. There are numerous types of cellular networks. An alternate solution network architecture, cell-free mMIMO, was considered. The installation of and connection to something like a CPU-central processing unit, as well recognised as an edge-cloud processing unit as well as CRAN data centre, will involve numerous dispersed single-antenna access points (APs). In order to collectively endorse the UEs through all the cohesion in joint transmission and reception, CPU controls the mechanism in N MIMO mode with really absence of cell borders. Cell-free mMIMO differs from standard network MIMO in that it operates with far more APs than UEs [6]. In the past, perfect CSI has been frequently assumed in performance analysis, so the incorporation of inadequate channel state information (CSI) was a significant analytical innovation. The study recommended using matched filtering or conjugate beamforming, also known as maximum ratio (MR) processing, locally at each AP while demonstrating that CPU processing that is partially or completely centralised can produce higher SE.

2 Existing Method

Both small cell systems and cellular networks are currently used approaches. We assume that in small cell systems, just one AP serves each user [7]. The accessible AP with the greatest average received useable signal power is chosen for each user. An AP becomes unavailable if another user has previously selected it. User by user,

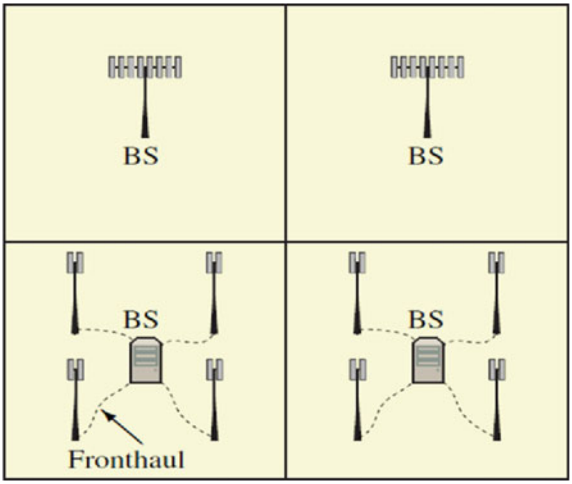


Fig. 1 Cellular connections with mMIMO BSs that have distributed or co-located arrays (top) (bottom)

in a random order, the APs are chosen. We take into account a time scale that is brief enough to prevent handovers between APs. Small cell systems prevent the channel from hardening [8] (Fig. 2).

As indicated in the picture, the cellular system has four cells inside a 1 km by one km space [9]. The uniform linear arrays are half-wavelength spaced on the multi-antenna APs, and Gaussian scattering local model with a fifteen degrees angular deviation is used to construct the spatial channel correlation (Fig. 3).

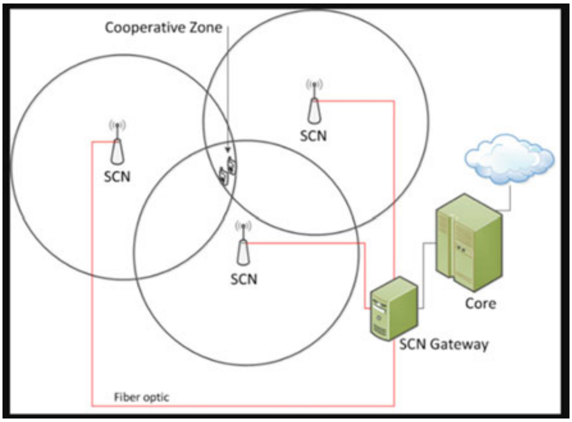


Fig. 2 Small cell network

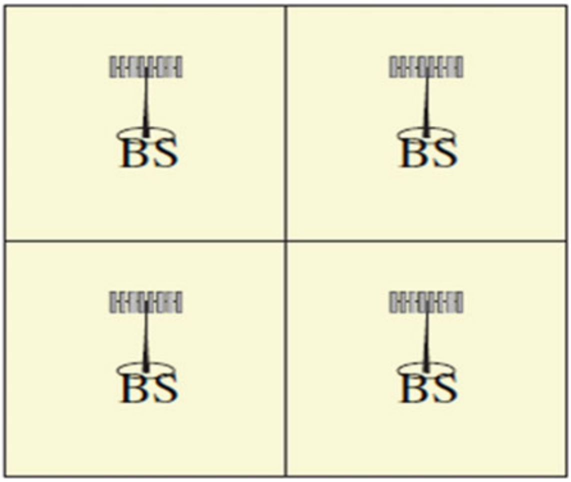


Fig. 3 Cellular network

3 Proposed System

The cell-free mMIMO network is made up of N antennas on each of L geographically dispersed APs. Uplink connections are employed to connect APs to a CPU, as shown in Fig. 4. Unlike wireless cellular networks, humans do not portion the network in to the cells or allocate users to particular base stations in a cell-free system [10]. Instead, designers suppose that what a province is covered by K distributed randomly single-antenna consumers, M distributed randomly smart antenna APs, and all these APs attached to a CPU through high-bandwidth links.

As an example, Figure depicts a cell-free system. Unlike a classical cell connection, each consumer in a cell absence scheme is assisted by multiple base stations. A

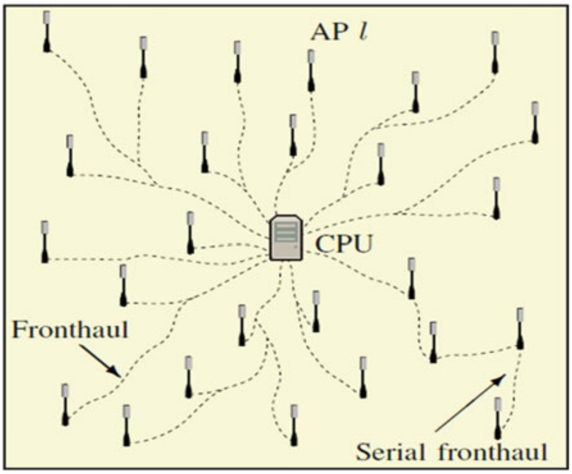


Fig. 4 Cell-free mMIMO network

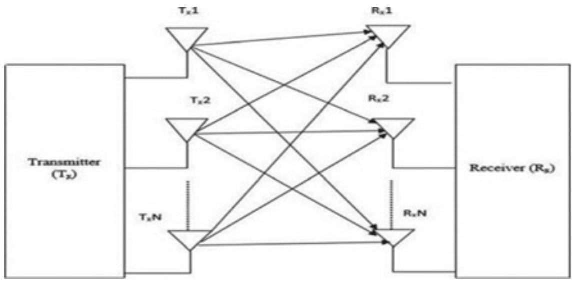


Fig. 5 Massive MIMO model

flat Rayleigh fading model is used for each OFDM SC [11]. The OFDM SC indicator has been misplaced for simplicity. The entire region is assumed to be small enough that the most significant signal propagation difference between two APs trying to reach a consumer is less than length of both the OFDM cyclical prefix, where β_{mk} is large-scale fading coefficient. This takes into consideration shadowing and path loss [12]. This coefficient is easily measured and tracked because it fluctuates slowly. We suppose that all these constants stand independent, i.i.d. arbitrary values that remain constant over the course of a coherent interval. For an OFDM wide-band system, β_{mk} is not frequency-dependent, while h_{mk} has frequency dependency and a Nyquist sampling interval that is proportional to the channel delay spread in frequency. We denote by

$$G \in C^{m\times k}, [G]_{mk} = g_{mk} \tag{1}$$

between all APs and users, the channel matrix. We also assume that the channel coefficients for uplink as well as downlink are the similar, or channel reciprocity. We concentrate on the case of users who move at less than 10 km/h. As this is often the case in real-world circumstances, we presume that the majority of our users are pedestrians (Fig. 5).

4 Massive MIMO Model

The $M \times 1$ received vector y at the BS is

$$y = \sqrt{p_u} Gx + n \tag{2}$$

where $n = \begin{bmatrix} n_1 \\ \vdots \\ n_m \end{bmatrix}$ is a vector of additive nil Gaussian mean noise.

Samples and the noise modification are set to 1, short of lost of generalisation. Let g_{mk} denotes the $M \times 1$ channel vector amongst BS and user K.

$$g_k = \begin{bmatrix} g_1 k \\ \vdots \\ g_m k \end{bmatrix} \tag{3}$$

$$E|g_{mk}|^2 = \beta_k$$

where β_k models the symmetrical attenuation besides shadow fading and it is a largest scale fading factor.

5 Channel Estimation

By considering the massive MIMO model, the channel estimation model is given by

$$y_{m\times k} = \sqrt{p_p} G_{m\times k} \emptyset_{K\times K} + N_{m\times k} \tag{4}$$

where $K =$ no. of pilot transmission.

The uplink transmission of data is shown in Fig. 6.

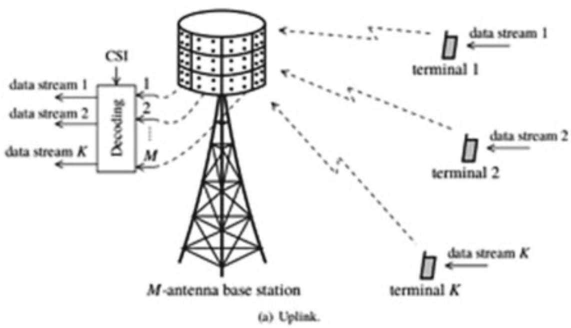
Now in the massive receiver, let us consider the user one as anticipated user, the conventional signal will be spilt as looked-for signal and also interference

$$y = \sqrt{p_u} g_{1x_1} + \sqrt{p_u} \sum_{i=2}^{k} g_{1x_i} \tag{5}$$

Therefore, the SINR can be obtained as

$$\text{SINR} = \frac{p_u \|g_1\|^2}{p_u \sum_{i=2}^{k} \beta_i + 1} \tag{6}$$

Fig. 6 Uplink transmission of data



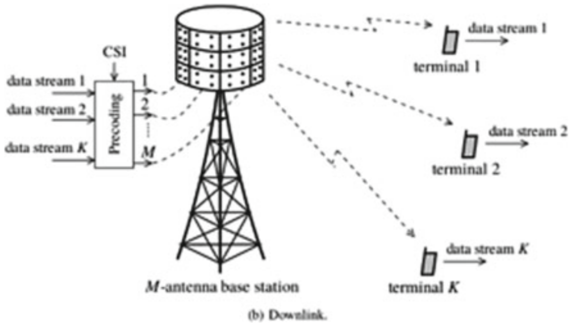


Fig. 7 Downlink transmission of data

where p_u is the power scaling, $p_u = \frac{E_u}{m}$ as the power of each user is decreased inversely as number of antenna by considering this power scaling factor.

The downlink transmission of data is shown in Fig. 7. As the massive MIMO operates in the TDD mode thus, channel estimate in the UL can be used in the DL. This is termed as channel reciprocity. Let W denotes the receiver combiner or beamformer.

The output of the beamformer is

$$\tilde{y} = [w_1^*, w_2^*, \ldots w_l^*] \begin{bmatrix} y_1 \\ \vdots \\ y_l \end{bmatrix} \quad (7)$$

6 Results

The estimated channel coefficients are shown in Fig. 8. This graphic most likely depicts the estimated channel coefficients obtained during the channel estimation procedure in massive MIMO systems. These coefficients indicate the characteristics of the channel between the base station and the users. The graph depicts the quality and accuracy of the calculated channel coefficients, which are critical for appropriate data transmission and reception in the system.

Figure 9, labelled estimated data, depicts the estimated data symbols obtained by channel estimation techniques. These estimated data symbols represent the transmitted data, which is reconstructed at the receiver using the estimated channel coefficients. The figure shows the estimating process's accuracy and the fidelity of the recovered data symbols.

Figure 10 depicts massive MIMO channel estimate under faulty CSI. When the channel state information is poor or contaminated, it most often illustrates the methodologies or algorithms used to estimate the channel properties in massive

Fig. 8 Estimated channel coefficients

G =

1.3942	0.8752	0.8007	1.2533	0.2255	1.4761	0.3497	0.6909	0.8679	0.8568
0.4001	0.4994	0.5293	1.6185	0.8145	0.8096	0.4210	0.9762	1.0774	1.7590
1.8640	0.8726	0.8878	0.9054	0.3444	0.2774	0.5866	1.7565	1.4293	0.6506
0.5422	0.5655	0.7973	0.9518	0.3930	0.8633	0.5927	0.3044	0.8960	0.5252
0.1641	0.7105	0.0693	1.2495	0.9188	0.7679	1.2799	0.8927	0.4982	0.6106
0.9972	1.2584	0.6483	1.0935	0.5980	1.1087	1.0121	0.8789	0.5119	0.3523
0.5130	1.5235	0.2957	1.3295	0.7567	1.6112	1.0190	0.4302	0.0257	1.4522
1.6004	1.0787	0.3726	1.5883	0.9197	1.3876	0.6322	1.6355	1.1074	0.4834
1.2077	0.6267	0.1631	1.0564	0.7332	0.4711	1.0652	0.4921	0.5167	0.7315
1.1172	1.5817	0.9124	0.6932	0.5330	1.2722	0.7756	0.3948	1.1482	0.5858
1.0145	0.9227	0.5760	1.7905	1.3374	1.9592	1.5308	1.2823	0.6613	0.7154
0.3947	1.8763	0.7936	1.4272	1.6322	0.3746	0.3464	0.7214	1.8298	1.2418
0.6127	1.1877	1.4701	0.4993	0.2329	1.1595	0.8263	0.4226	0.5169	0.3896
0.9922	1.8902	0.9690	0.6462	0.2414	0.9514	0.7528	1.2216	0.8295	1.2609
2.3674	1.2517	0.6255	0.3255	1.3801	0.9083	1.8102	1.1267	0.5098	0.7024
0.8959	1.0792	0.9213	1.1077	0.1241	0.6675	0.3456	0.2910	1.5311	0.6940
0.6170	1.9546	0.9853	0.7025	0.1041	1.0169	0.1391	0.9246	1.6749	0.3131

Fig. 9 Estimated data

Y =

-0.3746
0.5963
0.7674
-0.2091
-1.9349
-0.0699
0.7101
1.0317
-0.1147
-0.7343
-0.8676
0.6644
-1.0758
-0.1802
0.4689
-1.4508
0.2493

MIMO systems. This image depicts the channel estimate process visually and emphasises the problems and potential solutions for attaining accurate channel estimation despite incomplete CSI.

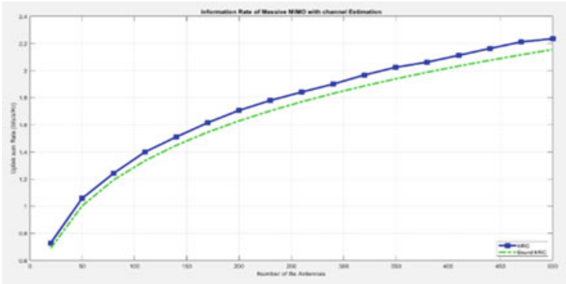


Fig. 10 Massive MIMO channel estimation

7 Conclusion

The study uses the MRC technique to analyse the efficacy of MIMO systems and obtains an expression for SINR. Three graphs illustrate the findings in the results section. Based on the data analysis, it is found that increasing the number of antennas enhances the performance of MIMO systems greatly. When the MIMO system has an appropriate number of antennas on both ends, it operates optimally. The performance is consistent whether sending from a base station to many mobile devices or receiving from multiple mobile devices. However, compared to systems with two or three antennas, using only one sending or receiving antenna leads in longer combining or broadcasting times. To maximise performance, it is advised to utilise an equal or even number of transmitting and receiving antennas, as configurations with different numbers of antennas may not produce optimal results. In future, research could concentrate on optimising antenna layouts and investigating advanced signal processing techniques to improve the performance of MIMO systems. Investigating the impact of various environmental parameters, including as channel conditions and interference levels, will also provide useful insights for real-world deployment scenarios. MIMO technology improvements have the potential to revolutionise wireless communication systems and accommodate the growing demand for high-speed and reliable data transfer in a variety of applications.

References

1. Bjornson E, Sanguinetti L (2019) Cell-free versus cellular massive MIMO: what processing is needed for cell-free to win? In: IEEE international workshop signal processing advances in wireless communication (SPAWC)
2. Vishnuvardhan D, Shaik K, Sai Bhavya P, Yadav P (2018) Monitoring physiological variables of mining workers. NCeTIMES#1.0, IJET ISSN: 2395-1303
3. Ngo HQ, Ashikhmin A, Yang H, Larsson EG, Marzetta TL (2017) Cell-free massive MIMO versus small cells. IEEE Trans Wireless Commun 16(3):1834–1850
4. Karimullah S, Vishnu Vardhan D (2022) Pin density technique for congestion estimation and reduction of optimized design during placement and routing. Appl Nanosci

5. Buzzi S, D'Andrea C (2017) Cell-free massive MIMO: user-centric approach. IEEE Commun Lett 6(6):706–709
6. Karimullah S, Vishnuvardhan D (2020) Experimental analysis of optimization techniques for placement and routing in ASIC design. In: ICDSMLA 2019, Lecture notes in electrical engineering 601, Springer Nature Singapore Pte Ltd
7. Aguerri IE, Zaidi A, Caire G, Shamai Shitz S (2019) On the capacity of cloud radio access networks with oblivious relaying. IEEE Trans Inf Theory 65(7):4575–4596
8. Karimullah S, Vishnu Vardhan D (2020) Iterative analysis of optimization algorithms for placement and routing in ASIC Design. In: ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
9. Van Chien T, Mollén C, Björnson E (2019) Large-scale-fading decoding in cellular massive MIMO systems with spatially correlated channels. IEEE Trans Commun 67(4):2746–2762
10. Karimullah S, Vishnu Vardhan D, Javeed Basha S (2020) Floorplanning for placement of modules in VLSI physical design using harmony search technique. In: ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
11. Aguerri IE, Zaidi A, Caire G, Shamai Shitz S (2019) On the capacity of cloud radio access networks with oblivious relaying. IEEE Trans Inf Theory 65(7):4575–4596
12. Karimullah S, Vishnuvardhan D (2018) A review paper on optimization of placement and routing techniques. At NC'e-TIMES # 1.0 IN, IJET. ISSN: 2395-1303

VIP Development of SPI Controller for Open-Power Processor-Based Fabless SOC

D. Anu Priya, D. Vishnuvardhan, G. Mamatha, Fahimuddin Shaik, and Shaik Karimullah

Abstract The SPI-Serial Peripheral Interface is considered as one of the most used bus protocols for attaching processors to associated devices with low/medium data transmission rates (SPI). In SoC applications, the SPI architecture is utilized to connect a large number of peripherals to the processor. The slave, which could be a sensor, monitor, or memory chip, is controlled by the master. A pre-packaged collection of code called Verification IP (VIP) is utilized for verification. A module designed to be used with a certain verification methodology, like UVM, may be the object at hand. Alternatively, it may be a set of statements for validating a bus protocol. The objective of the project is to provide Verification IP (VIP) blocks for an SPI controller coupled to an open POWER CPU, fabless SoC with an A2O core, utilizing the AXI4 interface. By creating test benches in UVM and System Verilog, the SPI controller's verification and VIP development are carried out. Software from Mentor Graphics® and Xilinx Vivado® was used for the simulation, synthesis, and verification, respectively.

Keywords QuetaSim · Functional verification · Reusable VIP · UVM—Universal Verification Methodology · SPI master core

D. Anu Priya
VLSI-SD, ECE Department, JNTUACE, Ananthapuramu, Andhra Pradesh, India

D. Vishnuvardhan
Department of Electronics and Communication Engineering, JNTUA, Anantapuramu, Andhra Pradesh, India
e-mail: vishnu.ece@jntua.ac.in

G. Mamatha
Department of ECE, JNTUA College of Engineering, Ananthapuramu, India
e-mail: gmamatha.ece@jntua.ac.in

F. Shaik (✉) · S. Karimullah
Department of ECE, Annamacharya Institute of Technology and Sciences, Rajampet, Kadapa, Andhra Pradesh, India
e-mail: fahimuddin.shaik.in@ieee.org

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_2

1 Introduction

Modern integrated circuits have developed quickly, which has raised the rigidness of IC design and made IC verification more difficult. Traditional verification approaches are no longer able to handle modern verification requirements, and 71–82% of the total design process is dedicated to verification prediction error expansion. System Verilog (SV), a unified circuit design as well as verification language, was introduced in 2002 by the Accellera Systems Initiative. The System Verilog language was created by combining constructions from various languages, including Vera, Super Log, C, VHDL, and Verilog [1]. Additionally, IEEE standardized (1800–2005) System Verilog in 2005. System Verilog supports behavioural, register transfer level, as well as gate-level descriptions. System Verilog facilitates testbench development with including OOC-object-oriented constructs, and cover groups, declarations, limited arbitrary constructs, but rather ASI-application-specific interface to the rest of languages [2]. The standardized UVM-Universal Verification Methodology is based on the Open Verification Methodology (OVM), which also combines several functionalities from the VMM-Verification Methodology Manual. The use of the UVM level increases verification productivity by providing a reusable verification system and verification components [3]. This work's verification study demonstrates the viability and utility of the discussed verification platform. SoC-System on Chip is a key determining that employs functional modules that are all interconnected on a single chip. Each SoC must be connected to other system parts effectively to provide faster, error-free communication for the system to be complete. It is crucial for core controller modules to communicate data with other external devices such external EEPROMs, DACs, and ADCs [4]. There are numerous types of communication protocols, including high speed protocols used for data transfers between entire systems, such as Ethernet, USB, SATA, and PCI-Express. Many people think of the Serial Peripheral Interface (SPI) as a lightweight communication mechanism [5]. The protocol's main benefit is that it may be used for communication between integrated circuits with on-board peripherals at low and medium data transmission speeds, where the considered serial bus offers a major price flexibility [6].

2 The Study's Objectives

The main purpose of this projected research project is for constructing a useful test bench which explains the AXI bus function model and SPI slave model-based SPI master controller. The following objectives help to attain the goal:

- To establish a correlation among the test bench variables and core controller, one must understand the SPI protocol architecture and the unique requirements of the AXI platform.
- Using cutting-edge verification methods like Coverage-Driven Functional Verification and Universal Verification Methodology.

- To create a recyclable Verification IP for an SPI main core that complies with AXI.

The primary contributions of this paper are as follows:

1. Learn about System Verilog, the Universal Verification Methodology, and the SPI sub-system architecture.
2. Creating a AXI bus model that serves as a connection between test circuit and the SPI DUT master device under test and the considered SPI slave model to enable closed-loop verification.
3. Design step by step testbench components with System Verilog frameworks, coverage, limited arbitrary stimuli, UVM libraries, but also assertions.
4. Confirm data transmission using various character widths and data types.

SPI can run at higher bandwidth and throughput than other protocols like UART and I2C. Typically, SPI Protocol enables communication between slave devices and the microcontroller on the host side. It is popular because it requires less control signals to operate. The specific SPI core that was the subject of this study performs the host side function of a slave AXI compliant device. The AXI Interface, Clock Generator, and Serial Shift Interface are the three main parts of the SPI Master Core Controller. Five 32-bit counters on the SPI primary controller can be modified using the AXI interface. It is considerable to implement a greater SPI Master/Slave somewhere around 900 and 1000 MHz. When two slaves are present, the core can be generated with much more adaptable handling of SPI-bus controls. By coding the control register of one important feature, the core that determines whether the SPI module is in master or slave mode. The SPI register file provides information about the current state of the data transfer process, including whether it has successfully completed or not, during operation [7]. The ability to adapt of SPI Interface IPs for different devices that used a parameterization approach is another crucial aspect. Communication issues between numerous devices are solved using TSM. The verification has gotten harder as a result of the current SoC's growing complexity [8]. In actuality, complicated SoC verification takes up 70% of the time required for product development.

The solution to the time to market problem is to reduce the verification work. Modern verification approaches are used to handle this increased complexity. In DFC, using constraint randomly processing technique is necessary for IP verification [9]. For this, a variety of tools are utilized, including coverage monitors and scoreboards. It is necessary to confirm that a data transfer protocol, alike the SPI data transfer protocol, adheres to the design requirements. Effective verification is achieved by using restricted random approach for improved functional coverage. EDA companies have been recommending more modern verification approaches and languages for many years. The effectiveness of any system-level verification technique and language depends on the created verification components' scalability and reusability [10]. The use of System Verilog with OOP is regarded as one of the best performing methods for greater level operation verification for modern sophisticated SOC systems. System Verilog offers a full verification environment, including assertion-based verification, direct and limited random generation, and metrics driven by coverage. The most recent functional verification approach is called the Universal Verification Methodology (UVM), and it makes use of System Verilog-coded base class libraries. UVM is based on earlier technique libraries like the AVM from Mentor, the OVM from Mentor and Cadence, the eRM from Verisity, and the VMM-RAL from Synopsys [11]. Users can now develop highly compatible, portable verification modules thanks to this standardization. Verification components are the name given to such modules. For complete systems, submodules, or protocols, they are encased and transformed into configurable verification environments that are ready to use. The basis for such applications is the extensive base class library. It provides coverage-driven restricted random verification, assertion-based verification, hardware acceleration, or emulation. It is simulation-oriented.

3 UVM Testbench

The designers simply applied an order of crucial stimulus elements to the circuit under test and compared the response to what was anticipated to complete the verification process. The design process was dominated by verification work as chip size continued to shrink and chip function grew more complex. Engineers use hardware description languages to articulate abstract frameworks of digital circuits as well as translate them in to the physical hardware. Validation is accomplished to use a testbench, a conceptual system that activates the inputs of a design being tested (DUT). According to functional verification, the design implementation adheres about the description. The content of the testbench enhances a reference architecture of the features and functions to be substantiated and relates the output of the model to the output of a design being tested. Functional verification's job is to check whether a design complies with requirements, not to prove it does [12]. Directed tests are used in the classical method of functional verification. Verification engineers design and apply a sequence of crucial stimuli directly to the test equipment, and then they determine whether the outcome is as was anticipated. Because the verification infrastructure can be put up with minimum effort, this method offers speedy first findings. But when a design gets more sophisticated, writing all the tests required to cover every aspect of it becomes a laborious and time-consuming effort [13]. Random events serve to hide the improbable circumstances and reveal bugs. However, the test environment has to automate the process for generating random stimuli, and in order to do so, a block that predicts, monitors results, and analyses them is required: a scoreboard. Additionally, a procedure known as functional coverage is employed to determine which instances of the random stimulus were covered as well as the design's current states. This type of test bench may take more time to construct, but random-based testing can really help with design verification by covering instances that cannot be covered by directed tests. Any HDL simulator that supports System Verilog should be able to use the UVM approach because it is an open-source, portable library developed by the Accellera Systems Initiative [14]. Additionally,

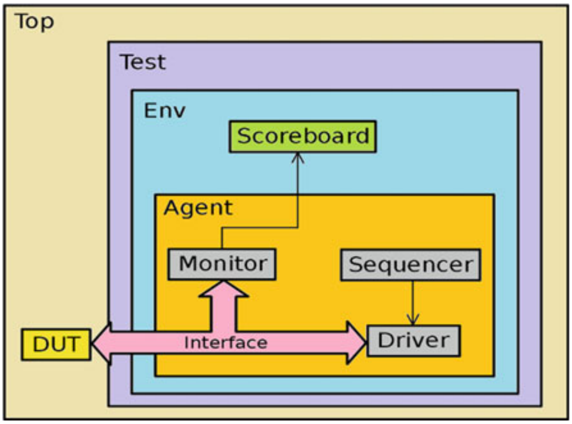


Fig. 1 UVM testbench architecture

UVM is derived with the help of the OVM library, which gives the methodology some context and experience. Reusability through the UVM API and standards for a common verification environment is important UVM features. Any verification engineer who is familiar with the approach can quickly modify the environment and comprehend it.

An approach called UVM was developed to create test platforms for design verification. For the purpose of designing modular testbenches for design verification, UVM has a specified process. Modular testbench components and stimulus can be designed and implemented with the aid of UVM's class library (Fig. 1).

UVM testbench top: Interfaces that connect the DUT to the testbench one and or even more implementations of architecture under test modules are typically included in the UVM testbench. Using Transaction-Level Modelling (TLM) interfaces, UVM offers channels of communication for transferring and receiving exchanges among the components. Because UVM Test do created dynamically during runtime, this same UVM testbench can be gathered once and used to run a variety of tests.

UVM test: UVM validation is indeed the high-level UVM element in UVM testbench. UVM Test normally performs important actions such as configuring principles in the config class as well as applying the proper stimulus towards the DUT by trying to call UVM sequence data throughout the platform. The basic test class creates and configures the high-level platform; successive tests of individual will enhance the base experiment by defining type of situation environment configurations like which extended scenes to execute, scope parameters, and so on.

UVM Environment: Verification components of UVM like agents, scoreboards, and even other environments are grouped together in the UVM environment, a container component class. All of the lower level verification components that are aimed at the DUT are contained in the high-level platform, which is a reconsiderable component. For the chosen setup, a variety of tests may be able to generate and send various types of traffic by instantiating the top-level environment class. The top-level environment's default settings can be overridden by UVM Test. Other child environments may also be instantiated by the master UVM environment. Each DUT interface may have a unique environment. UVM could be used, for instance, to develop reusable interface environments, including CPU environments, IP interface environments, PCIe environments, USB environments, and cluster environments.

UVM Agent: A class of container component is the UVM agent. Agent combines various verification components that deal with a certain DUT interface. Other parts of the agent include a sequencer, which controls the flow of stimuli, a driver, which delivers stimuli to the DUT input, and a monitor, which detects the DUT outputs. The driver receives the sequences that the sequencer has collected. At the DUT interface, the driver next translates a transaction sequence into a signal level. Agents have two different operating modes: active agent and passive agent. While passive agents can only sense the DUT active agents can produce stimuli. While the monitor only has a unidirectional interface.

Sequence item of UVM: The bottom object in the UVM priority is a UVM sequence item. AXI transaction, for instance, and it is used to construct sequences. The idea of a transaction was to separate the driver from data production while still handling bit-level activities using the DUT interface pins. Variables, restrictions, and even function calls that operate on themselves are all possible UVM sequence elements.

UVM Sequence: The selected verification environment must produce sequences utilizing the UVM sequence item after constructing it so that they can be transmitted for the sequencer. The combination of objects in an ordered sequence makes up a sequence. The necessity determines how the transactions are generated. The collection of values transmitted to the DUT is constrained or restricted by the sequence since the sequence item variables are often of the random kind. The simulation time being cut eventually helps.

UVM Driver: The variation-level methodology meets the bit, DUT clock, and pin actions in a component class known as a UVM Driver. Sequences from the sequencer are taken in by the driver, which translates them into bit-level operations before driving given data into the interface of DUT in accordance with the interface protocol. Driver drives DUT signals via the interface after receiving transaction items from the sequencer via the TLM port.

UVM Sequencer: The flow of sequence items between sequences generated and the driving component is controlled by the UVM sequencer. The UVM sequencer controls transaction flow from several sequences in a similar manner to an arbitrator. UVM sequencer and driver communicate using the TLM interface methods seq item export and seq item import, respectively.

UVM Monitor: The UVM monitor operates in the opposite manner from the UVM driver. To be delivered to the remaining UVM components, such as scoreboard, for analysis, monitor turns the DUT signal-level/bit-level values into transactions. The produced transactions are broadcast by the monitor using the analysis port. In order to maintain the UVM testbench's modularity, comparison with usually, the

UVM component that performs the intended output is the scoreboard. Using the monitor's analysis port, UVM monitor can also process post-converted transactions by gathering coverage information, recording, logging, inspecting, etc., or it can hand off the work to other components.

UVM Scoreboard: Checker feature is implemented on the UVM scoreboard. Typically, the checker compares the DUT response to an anticipated DUT response. Finally, the scoreboard contrasts actual data with data from the DUT output that was received. A reference model may be created using System Verilog, C, C++, or any other language. Reference models built in C, C++, and other programming languages can interact with the scoreboard thanks to the System Verilog Direct Programming Interface (System Verilog-DPI) API.

TLM connections: UVM offers implementation ports, export ports, and unidirectional ports for tying together various parts via TLM interfaces. When transferring data packets to numerous components or when the component to be connected is the lone class in the hierarchy, analysis ports are used.

4 Design of SPI

The SPI protocol uses full duplex mode for serial transmission. It has four wires that are used to send and receive data serially. SCLK-serial clock, MISO-master in slave out, MOSI-master out slave in, and SS are the four wires. SPI transmits and receives data bit by bit; we send and receive 8 bits. Since we are just using one slave in this instance, the slave select should always be low. When data must be transferred from the master to the slave, it is loaded into the master shift register and sent together with the clock signal. Through the MOSI line, the data bit will be sent to the slave's shift register. The slave shift register will get all 8 bits in a similar manner (Fig. 2).

Both the master and the slave use a straightforward RTL code. Data is concurrently shifted out serially, known as transmission, and shifted in serially, known as reception, during an SPI transfer process. Shifting of the data on the two serial data lines that are already present is synchronized by the serial clock (SCK) line. The slave select line, which is low, allows for individual slave device selection. Unselected slave devices do not obstruct SPI transmissions and reception [15]. When using the standard SPI protocol, the device needs to set the operating system to a frequency less than or equal to the slave device's maximum frequency. Typical communication frequencies range from 1 to 100 MHz. This same standard SPI protocol provides numerous devices

Fig. 2 SPI architecture

and just a single master. Because the chip-select line is low, the master broadcasts the appropriate chip-select bit to Logic 0 to select the slave device. When the MSTR bit in the control register is set to logic zero, the SPI will be setup as a slave. Serial clock (SCK) is the SPI clock input from the master in this mode, while master in slave out (MISO) and master out slave in (MOSI) will function as data output pins and a serial data output pin, respectively. As an input, the slave select (SS) pin will function. To start a transmission on the SPI, the SS pin must be asserted low, and it must stay low until the transmission is finished.

5 Verification of SPI

A sequence item is where the construction of the UVM environment starts. The class object known as a sequence item is typically an extension of the UVM transaction or UVM sequence item classes. All necessary data transfers are included, and utilizing UVM structures, they can be randomized or restricted to the designated border. Multiple sequence elements are generated by extending UVM sequences into sequences. To drive DUT pins, the produced sequences are fed into the driver. Various tasks make up the SPI master core driver. Getting the following sequence item is the driver's initial step. Second, we control the data transit. Third, we finish the sequence item by writing the packet to the UVM analysis port. A fork…join call is used to simultaneously execute the tasks. The creation of a monitor to watch the DUT's communication with the testbench is part of the testbench's design.

When the protocols are broken, it observes the pin-level transaction at the DUT's outputs and reports an error. All of these UVM parts are connected by the agent. In the monitor, the expected output of the DUT is forecasted, and the scoreboard contrasts the expected response with the actual response of the DUT. The env class is where the agent and scoreboard are created and connected. The SPI protocol and bus theory are both compatible with the SPI Master IP core. On the host side, the design is equivalent to wishbone bus specification-compliant slave devices. The AXI compliant SPI Master core device's general structure may be broken down into three functional components: the clock generator, the serial interface, and the AXI interface. The clk gen is responsible for generating this same clock signal from of the external entity clock wb clk I and generating the output signal s clk o based on the clock register bandwidth factors. Because there is no reveals that the relationship again for Serial Peripheral Interface, the clk gen module could indeed start generating a dependable serial clock transfer of information with an odd or perhaps even frequency segment in the register. The goal of the clock divider is to provide a frequency that is appropriate for the communication system. It is a crucial component of digital ASIC and FPGA architecture. It is possible to divide frequencies evenly in order to conserve resources. The wb_clk_i is divided by the core to produce the s_clk_o; any output clock frequency is attained by varying the value of the divider. The following is the expression of s_clk o and wb_clk_i.

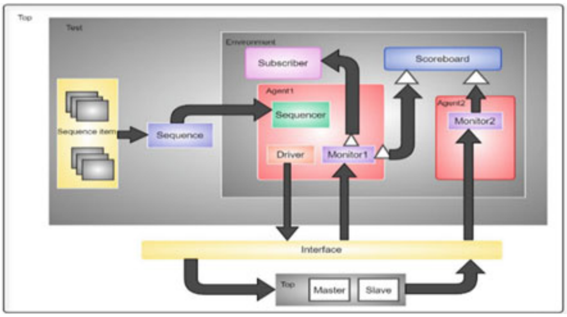


Fig. 3 SPI testbench architecture

$$f_{\text{sclk}} = f_{\text{wbclk}} / (\text{DIV IDER} + 1) * 2$$

Serial data transfer modules make up the information transfer core module. It is in charge of converting input MISO serial data in to the parallel out for MOSI transmission and input parallel in for MOSI transmission. The Actually obtain and transmit registers share the same flip-flops. If there was no start writing access to the data transmission register between transfers, it means that whatever data was did receive from of the input data connection in one transfer will be tried to communicate on the output data line in the next transfer. This has the advantage of using fewer hardware resources and thus consuming less power. While the master device transmits output data, the SPI Master core on the host side functions as a slave device to accept input data (Fig. 3).

6 Results

6.1 Transcript Results

See Fig. 4.

Fig. 4 Transcript results

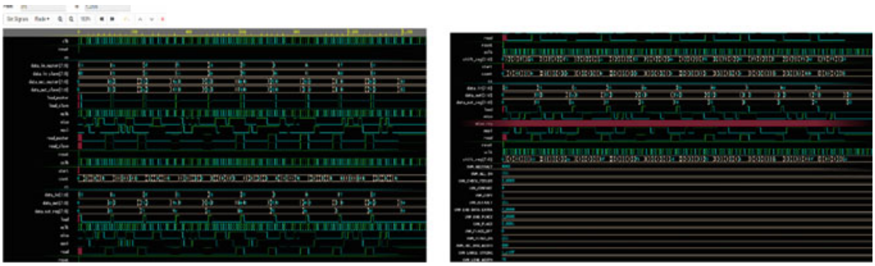


Fig. 5 SPI master to slave communication

6.2 Waveform Results

Using the AXI bus function model, the AXI and SPI master communicate with one another. With regard to the AXI protocol, the model primarily implements read, write, and reset functionalities. Master and slave communications are timed to the sclk_pad_clock, which is timed to the wb_clk base clock. Both the master and slave set up their control registers before the transfer begins. Flags like tx negedge/rx posedge, which specify the sampling edge of the send and receive signals, are found in the control register. Because SPI write and read input and output take place at the same only one buffer in a linear feedback shift method, the values of these three separate flags should be opposite. The divider register and slave select register are likewise set up by the master. The transfer is initiated when the go flag of the control signal is asserted after all SPI registers have been initialized. To synchronize the driver and monitor each endless loop component, the testbench leverages the flag transfer that is now in progress. Finally, as shown in Fig. 5, the transfer in progress signal is de-asserted after 32 clock cycles, signalling the end of transmission for the AXI interface to gather the data.

7 Conclusion and Future Scope

In this work, we develop an AXI compliant, reusable verification IP for the SPI Master core. We proposed a reusable testbench that includes a driver, monitor, SPI slave, scoreboard, agent, environment, coverage analysis, and assertions written using OOP using System Verilog and UVM. Additionally, to attain greater functional coverage, the approach of constrained randomization is employed. The outcomes of our simulation-based testing demonstrate the viability of the suggested VIP. This SPI Master core IP has been pre-validated and is prepared for SOC plug-and-play. This verification component can be applied to a variety of IP verification tasks throughout the project. Although Simulink verification is efficient for large systems, it does not provide an assured proof of correctness. Formal verification, on the other hand, provides a promised proof of accuracy by exhaustively having to cover the state space to uncover corner case bugs, but it does not scale well; at some point, verification becomes cumbersome due to the space vector explosion problem. Consider verification as just an extra step to simulation in future to increase the confidence of the verifiable system. The plan is to incorporate formal analysis in to the simulation settings.

- First-In-First-Out buffers can be added to the SPI master controller to accept data at various clock speeds.
- The AXI4 advanced specification can be added to the SPI master controller.
- The tests can be expanded to cover more SPI master controller variants in order to obtain 100% code coverage.

References

1. Abraham JA, Saab DG (2007) Tutorial T4A: formal verification techniques and tools for complex designs. In: 20th international conference on VLSI design (VLSID'07). p 6
2. Karimullah S, Vishnu Vardhan D (2022) Pin density technique for congestion estimation and reduction of optimized design during placement and routing. Appl Nanosci
3. Sutherland S, Davidmann S, Flake P (2010) System Verilog for design: a guide to using system Verilog for hardware design and modeling, 2nd edn. Springer Publishing Company, Incorporated
4. Karimullah S, Vishnuvardhan D (2020) Experimental analysis of optimization techniques for placement and routing in ASIC design. In: ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
5. IEEE Standard for Universal Verification Methodology language reference manual (2017)
6. Karimullah S, Vishnu Vardhan D, Basha SJ (2020) Floorplanning for placement of modules in VLSI physical design using harmony search technique. In: ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
7. Karimullah S, Vishnu Vardhan D (2020) Iterative analysis of optimization algorithms for placement and routing in ASIC design. In: ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
8. Zhou Z, Xie Z, Wang X, Wang T (2012) Development of verification environment for SPI master interface using system Verilog. In: IEEE 11th international conference on signal processing, vol 3. pp 2188–2192
9. Bromley J (2013) If system Verilog is so good, why do we need the UVM? In: Proceedings of the 2013 forum on specification and design languages (FDL). pp 1–7
10. Karimullah S, Vishnuvardhan D, Arif M, Gunjan VK, Shaik F, Noor-e-alam Siddiquee K (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Commun Mob Comput 2022:10, Article ID 3016709. https://doi.org/10.1155/2022/3016709
11. Ni W, Zhang J (2015) Research of reusability based on UVM verification. In: IEEE 11th international conference on ASIC (ASICON). pp 1–4
12. Fathy K, Salah K (2016) An efficient scenario based testing methodology using UVM. In: 17th international workshop on microprocessor and SOC test and verification (MTV). pp 57–60
13. Karimullah S, Basha SJ, Guruvyshnavi P, Sathish Kumar Reddy K, Navyatha B (2020) A genetic algorithm with fixed open approach for placements and routings. In: ICCCE. Publisher Springer, pp 599–610

14. Rajasekhar Reddy P, Sreekanth P, Arun Kumar K (2017) Serial peripheral interface-master universal verification component using UVM. Int J Adv Sci Technol Eng Manage Sci 3:27
15. Prasad R, Rani CS (2016) UART IP core verification by using UVM. In: IRF international conference

Cell-Free Massive MIMO Versus Small Cells

Neelagiri Sheela Rani, D. Vishnu Vardhan, G. Mamatha, Vinit Kumar Gunjan, and Fahimuddin Shaik

Abstract A huge number of access nodes are dispersed throughout the CFM-MIMO device and serve a great number of customer terminals within the equivalent time/frequency band. Cell frequencies are the focus of cell-free massive MIMO. However, mobile frequency bands may be used in conjunction with millimetre wave bands. There are no movable barriers or concepts of cells here. Multiple-entry, multiple-output (MIMO) equipment are used in a cell-free massive MIMO device. Distributed access points (APs) can act as a server a much relatively small number of customers using the same time/frequency investment based on carefully managed channel settings. Each of the most effectual APs and users has a single antenna. With the help of TDD-time-division duplex operating condition and the receipt of uplink received signal broadcast by clients, the APs learn about the channel's reputation. Multiplexing and de-multiplexing are accomplished by the APs using conjugate beam formation on the downlink but rather linked filtering on the uplink. Closed-form formulation is used to create max–min energy control algorithms for actual consumer downlink and uplink throughputs. Max–min electricity manipulate guarantees continually sturdy

N. Sheela Rani
DECS, ECEDepartment, JNTUA College of Engineering (Autonomous), Ananthapuramu, Andhra Pradesh, India

D. Vishnu Vardhan
JNTUA, Ananthapuramu, Andhra Pradesh, India
e-mail: vishnu.ece@jntua.ac.in

G. Mamatha
Department of ECE, JNTUA College of Engineering, Ananthapuramu, India
e-mail: gmamatha.ece@jntua.ac.in

V. K. Gunjan
Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, India
e-mail: vinit.gunjan@cmritonline.ac.in

F. Shaik (✉)
Department of ECE Annamacharya Institute of Technology and Science, Rajampet, Kadapa, Andhra Pradesh, India
e-mail: fahimuddin.shaik.in@ieee.org

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_3

service over the coverage area. While a pilot venture approach aids in lowering the consequences of pilot contamination, electricity manage is considerably greater vital.

Keywords Pilot · MIMO · Wi-Fi · Frequency · 5G · Wireless

1 Introduction

Multiple-input more than one-output generation, also known as MIMO, is the most effective realistic method for extensively enhancing spectral performance. It makes use of a couple of antennas to boom the throughput of Wi-Fi networks in order to triumph over the current electromagnetic spectrum limits [1]. In our research, we appoint M-MIMO, a prospective 5G wireless access era which could provide terrific throughput, dependability, and electricity efficiency with less signal processing. A base station with a couple of antennas is used to concurrently serve a massive wide variety of clients within the equal time–frequency aid [2]. The base stations have the option of installing massive antenna arrays in scattered or collocated configurations. Co-allocated M-MIMO structures, anywhere every overhaul antenna placed in a restricted area, obligate the benefit of requiring little backhaul. With the comparison, allocated huge MIMO buildings have service antennas dispersed over a wide area [3].

In this take a look at, we keep in mind a dispensed massive MIMO gadget, in which a good sized wide variety of get right of entry to factors, or carrier antennas, offer connectivity to a in moderation populated institution of impartial users dispersed over a good sized place [4].

Through a backhaul network, all get admission to factors collaborates phase coherently and use time-department duplex operation to offer the identical time–frequency useful resource to all clients. Cells and cell borders don't exist. Therefore, this generation is known as "cell-free massive MIMO." It is anticipated that these two architectures could greatly benefit CFM-MIMO since it associates the concepts of disseminated M-MIMO and only MIMO. Additionally, CFM-MIMO also can provide a maximum insurance opportunity because the customers are genuinely adjacent to the access points [5].

Although CFM-MIMO has a relevant processing unit, the data that may be sent between the access points and this CPU best is the payload information and the gradually changing power control coefficients. Instantaneous channel repute facts are not shared between the imperative unit and the get admission to points [6]. At the get admission to factors, uplink pilots estimate every channel. The downlinked records are pre-coded using the gotten channel estimations, and the uplinked facts are detected the usage of them s7]. To make sure that everyone network customers receive constantly high first-class provider, coherent processing throughout extensively scattered base station antennas is the intention of advanced backhaul. The unique characteristic of CFM-MIMO is how it functions: several single-antenna admittance points use computationally accurate (conjugate beam forming) signal processing to serve a significantly slighter number of customers immediately [8]. This makes it simpler to benefit from phenomena like high-quality propagation and channel hardening, which might be essential components of cell allowed massive MIMO [9].

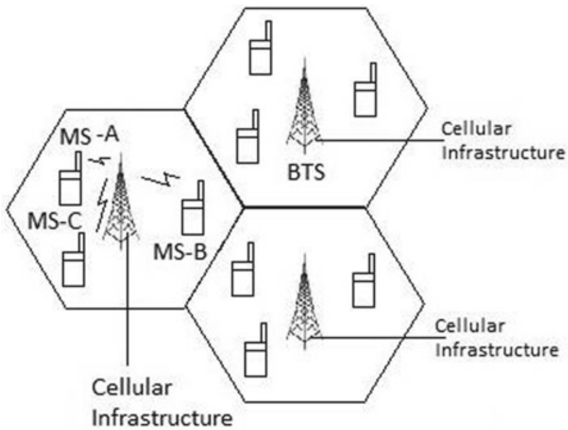
2 Exisiting System

A well-known mobile network every area is split into areas, or cells. Each of these cells has a base station (BS) deployed. One BS gives widespread coverage for numerous users. Backhaul connectivity is used to hyperlink the entire base stations to the middle community [10].

Figure 1 depicts cellular network. Base stations are related in a ramification of methods, consisting of famous person, mesh, and others. In the backhaul, they interface with the public switched data network, mobile switching centres, and public switched telephone network [11].

A cellular machine requires a reasonably complicated infrastructure, a block diagram shown in Fig. 2. A base station (BS) comprises a BTS and BSC [12]. Two sets of recommendations, the home location check in (HLR) and visitor location register (VLR) allow for international mobility and the usage of the same smartphone numbers. The equipment identity register (EIR) is a database that contains facts approximately the identities of mobile system.

Fig. 1 Cellular network



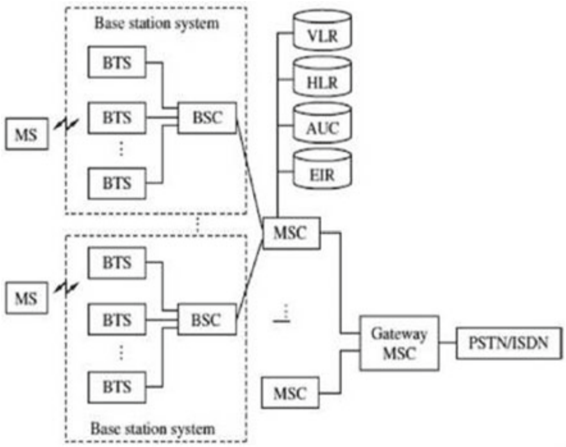


Fig. 2 Block diagram cellular network

3 Proposed System

In cell-free network, complete area is served via many access points (APs) and there is no idea of cells. Altogether, the APs are linked with CPU-central central processing unit. All the CPUs are interfaced with core network (CN). Cell-free network consists of antenna units, antenna processing units (APUs), central processing units (CPUs). Many APs serve all of the customers in their vicinity of have an effect on. One user is served by means of a couple of APs in mobile unfastened network [13]. A cell unfastened large MIMO machine with M access factors and K customers is taken into consideration. Each get right of entry to factor and consumer has a separate antenna, that is dispersed around an extensive place. Furthermore, a backhaul community is used to attach all get admission to factors to an important processing unit shown in Fig. 3.

Here, g_{mk} denotes the medium constant amongst kth customer and mth access point. The medium's g_{mk} is demonstrated as:

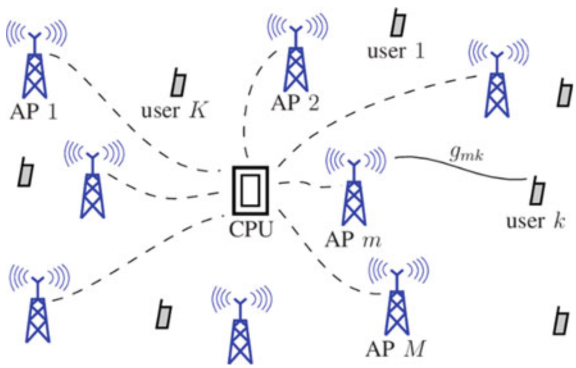


Fig. 3 Cell-free massive MIMO system

$$g_{mk} = \beta_{mk}^{\left(\frac{1}{2}\right)} * h_{mk} \tag{1}$$

Here, h_{mk} characterizes the low level fading along with β_{mk} characterizes the high-level fading. Hence to undertake h_{mk}, $m = 1, 2, \ldots, M$, $K = 1, \ldots K$, are autonomous along with identically disseminated. Uplink schooling, downlink payload statistics switch, and uplink payload information transmission make up every coherence interval [14].

A **Uplink Training:**

The values of g_{mk} and h_{mk} depend on frequency but m_k possesses a broad spectral bandwidth and is frequency independent.

Over a coherence interval of time and an intensity coherence interval, the propagation streams are assumed to be piecewise constant. Let τ represents the coherence duration length, which is the result of coherence period and coherence bandwidth, as well as τcf represents the static loading time frames, where cf originated.

B **Downlink Payload Data Transmission:**

The entry points use coupled beam establishment to send a notification to the K users and behave the channel forecasts as the actual channel.

The base station transmits the data symbols x_k to the users $k = 1, 2, \ldots, K$. The transmitted signal $s(t)$ can be represented as:

$$s(t) = \sum [h_k * x_k] * p(t) \tag{2}$$

The received signal $y_{k(t)}$ at user k can be expressed as:

$$y_{k(t)} = h_k^H * \left[\sum\left[h_j * x_j\right] * p(t)\right] + n_{k(t)} \tag{3}$$

where

- K: The total number of users.
- N: The total number of transmit antennas at the base station.
- h_k: The channel vector between the base station and user k, with dimension $N \times 1$.
- x_k: The data symbol vector intended for user k, with dimension 1×1.
- n_k: The additive white Gaussian noise vector at user k, with dimension N × 1.

4 Pilot Assignment and Power Control

Power control and pilot project can be accomplished individualistically for the reason that the pilots aren't power precise.

(a) **GPA-Greedy Pilot Assignment:**

Due to the short coherence length, various clients frequently need to hire non-orthogonal pilot sequences. There are τ^{cf} perpendicular pilot arrangements if the length of the pilot categorizations is τ^{cf}. Here, the case that $\tau^{\mathrm{cf}} < K$ is the focus. If $\tau^{\mathrm{cf}} > K$, we just give the K users K orthogonal pilot sequences.

(b) **Power Control:**

Although power management is often a well-researched area, the maximum to minimum difficulties in power control that ascend whenever cell unrestricted M-MIMO optimized is unquestionably original. Energy manipulation is done within the CPU. We discover the coefficients of power control that exploit the minimal compared to the downlink amounts for every user in downlink given realizations fading.

5 Small-Cell System

In small-cellular systems, each consumer is served by means of best one AP. Each person chooses the accessible get admission to factor with the very best common useable sign energy obtained. This get right of entry to factor is currently unavailable because any other consumer previously chose it. Because every other user formerly decided on this gets right of entry to factor, it now becomes unavailable. The access factor is selected randomly and consumer with the aid of consumer order.

The channel does no longer harden in tiny cellular systems, compared to cellular-huge MIMO. In the tiny cell case, the operative channel is an only one fading Rayleigh particle coefficient, so MIMO is an additive manufacturing technology of M-vectors and thus rounds its propose [15]. Since both consumers and APs must recognize their strong channel advantage that enables demodulation of the symbols, uplink and downlink training are necessary.

6 Results

A Uplink Results

The variable declaration window is depicted in Fig. 4. This illustration most likely depicts the variables or parameters used in the analysis or simulation of cell-free massive MIMO and small cells. Variables such as the number of antennas, transmit

Fig. 4 Variable declaration window

power, channel conditions, user locations, and any other relevant parameters that affect system performance could be included.

Figure 5, titled "Cumulative distribution versus per user Uplink Net throughput (Mbits/sec)," depicts the cumulative distribution of uplink net throughput in cell-free massive MIMO or small-cell systems. The x-axis is most likely the uplink net throughput in megabits per second (Mbps), and the y-axis is the cumulative distribution function (CDF). The graph allows you to compare the performance of the two systems in terms of user throughput.

B **Downlink**

Figure 6 depicts the input declaration window for the downlink scenario. This figure most likely represents the input parameters or variables for the downlink performance analysis or simulation. The inputs value of 50 indicates that the downlink analysis takes into account 50 input variables or parameters. User density, signal-to-interference-plus-noise ratio (SINR), path loss, and any other relevant parameters could be included as inputs.

Figure 7, titled "Cumulative distribution versus Net throughput of each user downlink (Mbits/sec)," depicts the cumulative distribution of net throughput for each user in cell-free massive MIMO or small cell systems. The x-axis represents net throughput in Mbps, and the y-axis represents the cumulative distribution function, as in Fig. 5. This graph compares the performance of the two systems in terms of the throughput achieved by users in the downlink scenario.

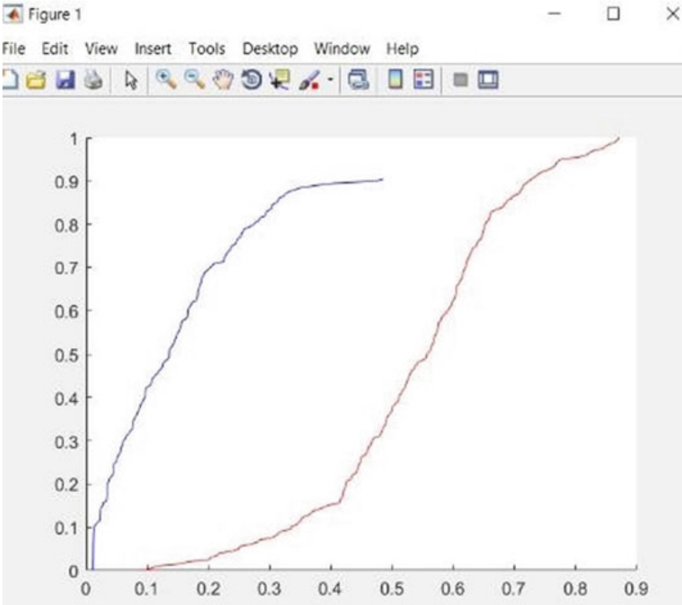


Fig. 5 Cumulative distribution versus per user uplink net throughput (Mbits/sec)

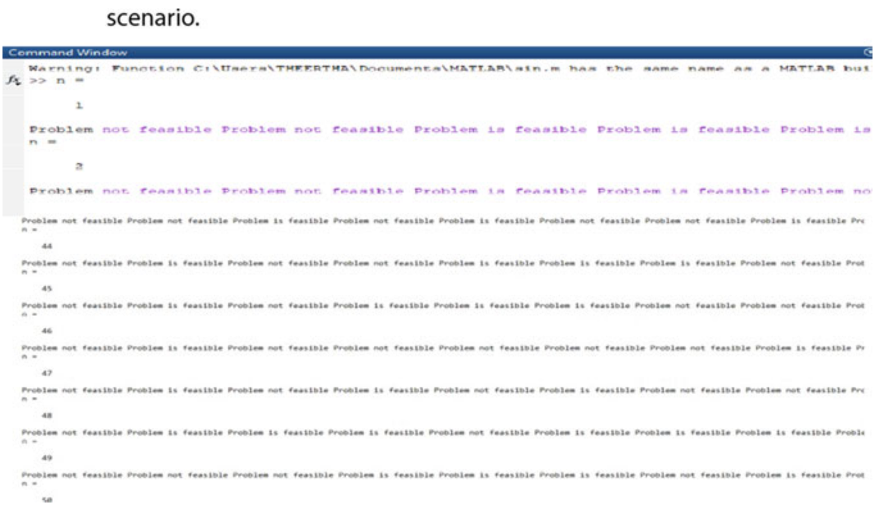


Fig. 6 Input declaration window where inputs = 50

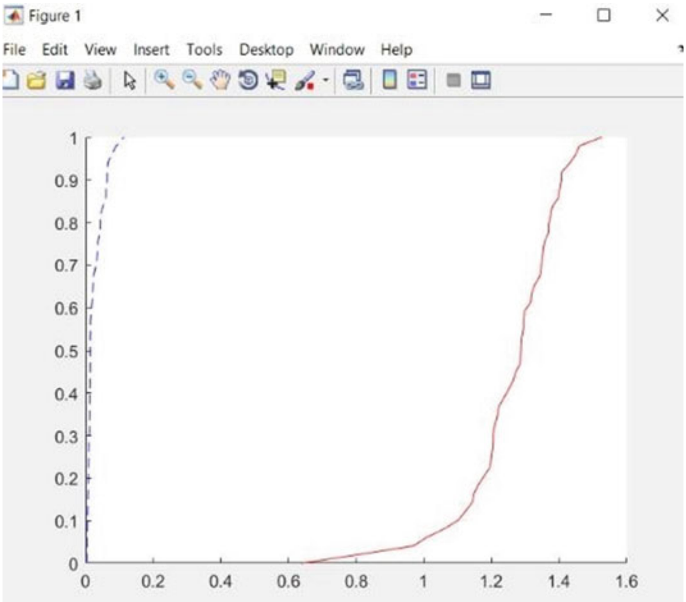


Fig. 7 Cumulative distribution versus net throughput of each user downlink (Mbits/sec)

7 Conclusion and Future Scope

Research results show that cell-free massive MIMO systems outperform small-cellular structures in terms of throughput. Cell-free systems, in particular, display more resilience to shadow fading correlation than small-cellular structures. Cell-free massive MIMO with shadowing correlation achieves much greater probabilities of 95% per person throughputs than small-cell systems, according to the research. In addition, when compared to cell-free massive MIMO, the implementation complexity of small-cellular structures is significantly lower. Cell-free and user-centric architectures appear to offer numerous network MIMO benefits while needing less backhaul. These findings point to cell-free massive MIMO as a possible method for improving wireless transmission performance and providing improved coverage in sixth-generation (6G) wireless communication systems.

Future research on 6G wireless communication systems will be able to analyse and investigate the capabilities and limitations of cell-free massive MIMO in various circumstances. Further research can concentrate on optimizing resource allocation algorithms, reducing interference, and increasing energy efficiency in cell-free massive MIMO systems. Furthermore, research into the integration of other developing technologies, including as artificial intelligence and machine learning, with cell-free massive MIMO can provide vital insight into the possibilities of these combined approaches. These initiatives are aimed at improving the performance and capacities of wireless communication systems, which will lead to more efficient and dependable wireless networks in future.

References

1. Ngo HQ, Ashikhmin A, Yang H, Larsson EG, Marzetta TL (2015) Cell-free massive Mimo: uniformly great service for everyone. In: Proceedings of the IEEE international workshop on signal processing advances in wireless communications (SPAWC), Stockholm, Sweden
2. Karimullah S, Basha SJ, Guruvyshnavi P, Sathish Kumar Reddy K, Navyatha B (2020) A genetic algorithm with fixed open approach for placements and routings. In: ICCCE. Publisher Springer, pp 599–610
3. Karimullah S, Vishnu Vardhan D (2020) Iterative analysis of optimization algorithms for placement and routing in Asic design. In: ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
4. Yang H, Marzetta TL (2014) A macro cellular wireless network with uniformly high user throughputs. In: Proceedings of the IEEE vehicular technology conference (VTC)
5. Karimullah S, Vishnu Vardhan D (2022) Pin density technique for congestion estimation and reduction of optimized design during placement and routing. Appl Nanosci
6. Matthaiou M, Zhong C, Mckay MR, Ratnarajah T (2013) Sum rate analysis of Zf receivers in distributed Mimo systems. IEEE J Sel Areas Commun 31(2):180–191
7. Karimullah S, Vishnuvardhan D (2020) Experimental analysis of optimization techniques for placement and routing in ASIC design. In: ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd
8. Arzetta Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wireless Commun 9(11):3590–3600
9. Vishnuvardhan D, Karimullah S (2018) Monitering physiological variables of mining workers. Ncetimes#1.0, IJET ISSN:2395-1303
10. Zhou S, Zhao M, Xu X, Wang J, Yao Y (2003) Distributed wireless communication system: a new architecture for future public wireless access. IEEE Commun Mag 41(3):108–113
11. Marzetta TL, Larsson EG, Yang H, Ngo HQ (2016) Fundamentals of massive Mimo. Cambridge University Press, Cambridge, UK
12. Ngo HIQ, Ashikhmin A, Yang H, Larsson EG, Marzetta TL (2017) IEEE Trans Wireless Commun
13. Vishnuvardhan D, Karimullah S, Sai Bhavya P, Yadav P (2018) Monitoring physiological variables of mining workers. Ncetimes#1.0, IJET ISSN:2395-1303
14. Jeevitheswari B, Karimullah S, Vishnuvardhan D (2017) A low power DFT based CMOS receiver with dual use of power lines. IJSTR 6(22):4470–4475. ISSN 2319-8885
15. Karimullah S, Vishnu Vardhan D, Basha SJ (2020) Floorplanning for placement of modules in VLSI physical design using harmony search technique. ICDSMLA 2019, Lecture notes in electrical engineering 601. Springer Nature Singapore Pte Ltd

High-Precision Navigation Using Particle Swarm Optimization-Based KF

Nalineekumari Arasavali, Sivakumari Gorle, Sasibhushanarao Gottapu, and N. Ashok Kumar

Abstract Position accuracy measures a system's capability to provide accurate estimates and is dependent mainly on the type of algorithm and the measurement noise. Kalman filter estimator (KFE) is the most common positioning algorithm used in real-time applications because of its atypical performance in a variety of different applications. The prior knowledge of parameters decides the accuracy of navigational solution in tracking applications. The deciding parameters are model uncertainty, measurement noise and estimation error. This paper presents a new navigational algorithm, particle swarm optimized Kalman filter (PSOKF). Particle swarm optimization is used in finding optimal Kalman filter parameters by minimizing the objective function, which incorporates factors like variance of measurement noise, variance of process noise and covariance of initial state estimation. The positioning performance of PSOKF is illustrated with data collected in the area of southern region of Indian subcontinent and GPS receiver is located at Lat/Lon: 17:72oN/83.32oE.

Keywords GPS · Kalman filter · PSO

1 Introduction

The GPS receiver measurement accuracy may be effected by many different error sources such as software errors, or atmospheric conditions on the day of the measurement. GPS receiver's location is the reference location. This receiver compares the calculated location to the known location by using satellite signals [1]. In order to determine where the receiver is located accurately, precise range measurements are needed. The measurement includes arrival time of the signal, the delay caused by the ionosphere and troposphere, as well as measurement noise and multipath

N. Arasavali (✉) · S. Gorle · S. Gottapu · N. Ashok Kumar
Dadi Institute of Engineering and Technology, Anakapalle, India
e-mail: naliniarasavali@gmail.com

S. Gorle
e-mail: gsivakumari@diet.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_4

errors. Errors related to troposphere and ionosphere can be corrected, but remaining errors may cause inaccuracies in the user's position. Many correction models for ionospheric and tropospheric error corrections are available today [2]. The error may cause with chosen navigational solution. The existing navigational solutions are least squares [3], Kalman filter, extended Kalman filter, unscented Kalman filter, particle filter, etc. All these algorithms are efficient but their performance depends on selection of initial parameters applied in algorithms, respectively. The initial parameters selection while estimating position uncertainty [4] by trial and error method take lot of computational time and suitable for real-time applications. The main objective of this work is to select initial parameters of Kalman filter algorithm using swarm intelligence technique. The most efficient particle swarm optimization algorithm is used for tuning of Kalman parameters and developed a new navigational solution PSOKF.

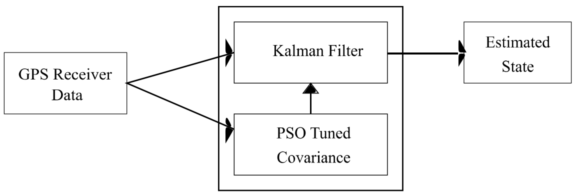
2 Methods

The proposed methodology consists three important steps. They are (1) Kalman filter application for data collected from GPS receiver located at southern region of Indian subcontinent. (2) Tuning of parameters using particle swarm optimization. (3) Incorporation of optimal parameters into Kalman filter algorithm. The process of proposed model PSOKF is depicted in Fig. 1.

2.1 Kalman Filter

Kalman filtering is an entrenched recursive algorithm for nonlinear problems. The concept of recursive computation involved in Kalman filter has been the subject of extensive research and implementation, especially in the field of autonomous or aided navigation. Many applications of this filter can be found primarily in the military and space sectors. Precise position estimation is the big challenge for any existing or upcoming navigation solutions. Kalman filters have been used for many years as a perfect solution to tracking and data prediction problems. The values predicted by Kalman filter are not true values but closer to true values, the difference is due to measurement noise and process noise in the system [5]. Thus uncertainty

Fig. 1 Block diagram of PSOKF



is measured by algorithm and then KF finds the weighted average of prediction and measurement to estimate accurate values. Values with the least uncertainty are given the most weight. In this way, estimates obtained from the method are more accurate. Prediction and correction are two steps involved in Kalman filter [6] and it involves mathematical formulation from Eqs. (1) to (7).

Predicted state estimate:

$$s_{\xi/\xi-1} = \varphi_{\text{tr}} s_{\xi-1/\xi-1} \tag{1}$$

Covariance of predicted estimate:

$$p_{\xi/\xi-1} = \varphi_{\text{tr}} P_0 \varphi_{\text{tr}}^T + Q_v \tag{2}$$

The corresponding time update statements provide covariance of estimated state from time step $\xi - 1$ to step ξ.

Measurement residual

$$y_{\xi/\xi} = y_{\xi/\xi-1} - H s_{\xi/\xi} \tag{3}$$

Residual covariance

$$p_{\xi/\xi} = H_\xi p_{\xi/\xi-1} H_\xi^T + R_v \tag{4}$$

The gain of KF is given as in Eq. (5)

$$k = p_{\xi/\xi-1} H_\xi^T (H_\xi p_{\xi/\xi} H_\xi^T + R_v)^{-1} \tag{5}$$

Updated state estimate

$$x_{\xi/\xi} = x_{\xi/\xi-1} + k y_{\xi/\xi} \tag{6}$$

Updated covariance estimate

$$p_{\xi/\xi} = (1 - k H_\xi) p_{\xi/\xi-1} \tag{7}$$

This statement involves: H as observation matrix; k, as Kalman gain; P_0 as covariance of initial state estimation; Q_v as variance of process noise; and R_v as variance of measurement noise covariance. R_v is considered as $E[ee^T]$. This Kalman filtering technique applied to process data to find measurements with assumed values of covariance of state and variances of measurement and process noises.

2.2 Particle Swarm Optimization

The position uncertainty is estimated by evolutionary algorithms, and the research is done with several algorithms [7]. In past several years, particle swarm optimization (PSO) proved as a successful random optimization technique, inspired by social behavior of group of birds, in many research areas [8]. It is observed that PSO is the fastest and provide better results compared with other methods. Unlike GA, PSO has no crossover and mutation parameters. In PSO, individual solution is known as particle, set of vectors described as swarm. The particles change their components and move in a space. Each particle is characterized by its position and velocity vectors [9]. Each particle has its individual personal best and social global best values, where global best is the personal pest of neighbor [10]. At each time step, velocity needs to be updated to achieve best solution values. PSO has velocity update and position update Eqs. (8) and (9) which are given below.

Velocity update:

$$\begin{aligned} \mathrm{vel}_i(\xi+1) = w \times \mathrm{vel}_i(\xi) + c1 \times \mathrm{rand} \times (pb_i(\xi) - x_i(\xi)) \\ + c2 \times \mathrm{rand} \times (gb(\xi) - x_i(\xi)) \end{aligned} \tag{8}$$

Position update:

$$x_i(\xi+1) = x_i(\xi) + v_i(\xi+1) \tag{9}$$

where w = Inertia weight (0.5)

$c1$ and $c2$ are coefficients of personal and global learning. In this paper, the objective function is chosen for PSO is as given Eq. (10). Steps following in PSO are described in given below Algorithm 1.

$$F = \frac{\|P_0\|}{N-1} + \frac{\|R_v\|}{N-1} + \frac{\|Q_v\|}{N-1} \tag{10}$$

Algorithm 1 Pseudocode for PSO

```
Input: Problem size, Size of Population
Output: P_gbest
Population = φ
P_gbest = φ
For (i = 1 to Size of Population)
            P_velocity = Randomvelocity()
            P_position = Randomposition(Size of POP)
            P_best = pposition
     If (obj(p_best) ≤ obj(p_gbest))
        P_gbest = p_pbest
     End
For(p ε population)
            P_velocity = Updatevelocity(p_velocity, p_pbest, p_gbest)
            P_position = Updateposition(p_velocity, p_position)
     If (obj(p_position) ≤ obj(p_pbest))
            P_pbest = p_position
     If (obj(p_pbest) ≤ obj(p_gbest)
            P_gbest = p_pbest
     End
     End
End
End
Return (p_gbest)
End
```

2.3 Optimal Parameters into Kalman Filter

Parameter values obtained through PSO technique are taken into the expressions of KF. The processed data through navigation solution given as input to PSO. Based on this information, PSO gives the optimal values for P_0, Q_v and R_v by that position accuracy is observed. The process is shown in the block diagram Fig. 1.

3 Results

As a result, estimation error values in ECEF coordinates are reduced by approximately 9.88 m, 17.44 m, 5.80 m, respectively, compared to conventional KF. The X_e, Y_e and Z_e error values are given in Table 1 from the processed data collected at Andhra University located in Visakhapatnam during 24 h on 3rd March 2011. Table 1 gives the performance comparison of both PSO-based KF technique and estimation algorithm, KF. During 10:00 h on March 3, 2011 at Visakhapatnam, the position

Table 1 Error values for KF with PSO

SOW	*X*-direction	*Y*-direction	*Z*-direction
439,200	46.0069	73.3345	25.9727
446,400	45.7881	93.9534	31.8138
453,600	25.4579	32.8975	12.3316
460,800	42.6756	74.5008	23.9273
468,000	39.2435	78.9782	32.9825
475,200	44.6843	65.0481	27.3054
482,400	45.7688	97.8954	31.09325
489,600	40.6701	77.9500	25.4142
496,800	50.1696	91.2232	30.8904
504,000	47.6162	82.6082	22.3413
511,200	37.3357	79.1713	30.6553

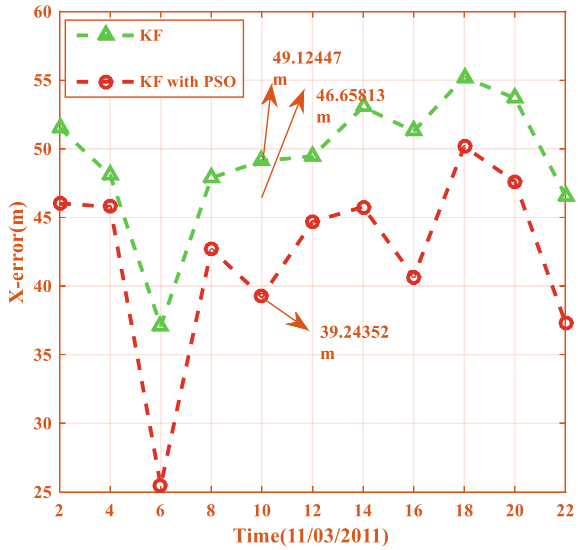


Fig. 2 Positioning errors in *X*-direction

error values are 39.245, 78.9782 and 32.9825 m in *X*-, *Y*- and *Z*-directions, respectively. PSO is more efficient than genetic algorithm. PSO algorithm's convergence rate is more than genetic algorithm. Figure 2 depicts how positioning performance improved in *X*-direction when compared to conventional KF. At 6th hour, the difference is more and it is very close during 4th hour. At every epoch of time period, error difference is considerable while estimating the position.

Figure 3 shows comparison of KF and PSOKF in *Y*-direction. Figure 4 shows clarity to consider PSOKF as prominent solution in precise positioning applications.

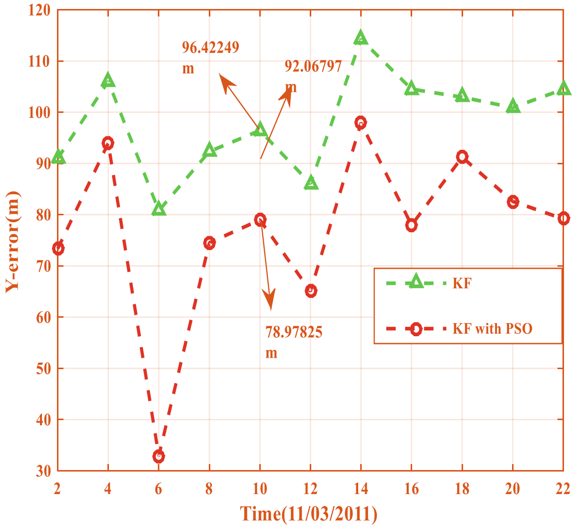


Fig. 3 Positioning errors in Y-direction

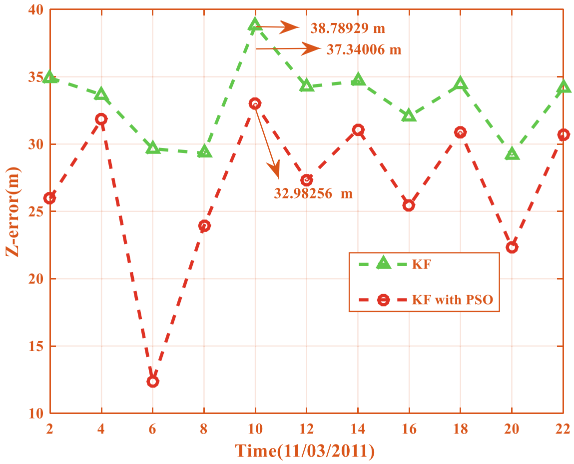


Fig. 4 Positioning difference in Z-direction

4 Discussions

The achievement of proposed algorithm is compared in terms of positioning accuracy with standard kalman filter with the data form GPS receiver is at Lat/Lon:17.72°N/ 83.32°E. At 12th hour, it is measured that position error values with KF are 49.45 m, 86.06 and 34.25 m in X-, Y- and Z-directions. The percentage of improvement in accuracy values when the comparison is done between PSOKF and KF are 10.67, 32.31 and 25.45%. It is observed that more accuracy is achieved in Y-direction. The developed PSOKF algorithm showing better performance than traditional navigational solutions.

5 Conclusion and Future Work

This work presented a new navigational solution PSOKF and it is observed that PSOKF is an efficient algorithm for solving positioning issues in tracking applications. The trial and error method increases the time of computation while assuming the initial parameters of existing widely used KF algorithm. In order to provide best positioning values optimal values of parameters required. The optimization is done with most efficient swarm intelligence technique particle swarm optimization. And efficiency of PSOKF is verified experimentally. Around 33% accuracy is achieved in Y-direction. It is concluded that proposed PSOKF is a better technique for all issues existing in real-time precise positioning fields undoubtedly. The future work involves the comparison of invented PSOKF with existing EKF, unscented KF and particle filter, etc.

References

1. Ashok Kumar N, Sasibhushana Rao G, Lavanya B (2018) Novel BAT algorithm for position estimation of a GPS receiver located in coastal region of Southern India. Procedia Comput Sci 143:860–867. ISSN 1877-0509. https://doi.org/10.1016/j.procs.2018.10.369
2. Bhushana Rao GS (2008) Ionospheric delay estimation for improving the global positioning system position accuracy. IETE J Res 54(1):23–29
3. Kumar PS, Dutt VSI (2021) Interpretation of statistical errors for precise GPS receiver positioning using recursive LS and EKF approaches. In: Communication software and networks. Springer, Singapore, pp 415–423
4. Rao GS, Siripurapu D, Bagadi L (2018) Elevation and position uncertainty based KF model for position accuracy improvement. Procedia Comput Sci 143:914–920
5. Zhang Q, Zhao L, Zhao L, Zhou J (2018) An improved robust adaptive Kalman filter for GNSS precise point positioning. IEEE Sensors J 18(10):4176–4186. https://doi.org/10.1109/JSEN.2018.2820097
6. Laveti G, Sasibhushana Rao G, Bidikar B (2016) Modified Kalman filter for GPS position estimation over the Indian sub continent. Procedia Comput Sci 87: 198–203. ISSN 1877–0509. https://doi.org/10.1016/j.procs.2016.05.148
7. Rao GS, Bagadi L (2018) Hybrid technique for GPS receiver position applications. Procedia Comput Sci 143:899–906
8. Tharwat A, Schenck W (2021) A conceptual and practical comparison of PSO-style optimization algorithms. Expert Syst Appl 167:114430. ISSN 0957-4174. https://doi.org/10.1016/j.eswa.2020.114430
9. Dereli S, Köker R (2021) Strengthening the PSO algorithm with a new technique inspired by the golf game and solving the complex engineering problem. Complex Intell Syst 7(3):1515–1526. https://doi.org/10.1007/s40747-021-00292-2
10. Tsai S-B, Wu Y, Song Q (2021) Improved particle swarm optimization algorithm in power system network reconfiguration. Math Prob Eng. 5574501:1024-123X. https://doi.org/10.1155/2021/5574501

Recent Advancements for Detection and Prediction of Breast Cancer Using Deep Learning: A Review?

P. Jyothi and Srinivas Koppu

Abstract The mortality rate of breast cancer the world over is very high. The probability of breast cancer in women keeps increasing with age. The cause of breast cancer can be from heredity to excessive breastfeeding. In this article, we have analyzed the existing imaging techniques available for the detection of breast cancer. The techniques of X-ray mammography, ultrasound imaging, and magnetic resonance imaging are analyzed for their advantages and disadvantages with respect to the detection of small cancerous regions, frequency, exposure to radiation, and providing comfort to patients. The deep learning and machine learning techniques are analyzed. The article is summed up with survey the detection and different approaches are used to predict breast cancer like deep learning, machine learning, and mammography.

Keywords Breast cancer · Neural networks · Machine learning · CNN · Deep learning

1 Introduction

According to WHO data, globally 2.3 million women were diagnosed with breast cancer and more than 685,000 lost their lives. Breast cancer is the most prevalent of all types of cancer. Women of any age group and country can get breast cancer after attaining puberty. The probability is high with an increase in age. There has not been any significant change in the death rate due to breast cancer from the year 1930 to 1970 [1]. After 1970, many countries started the cancer early detection program. The aim of such programs was to find a proper solution to the disease. The reasons for breast cancer are the use of alcohol, obesity, getting older, exposure to harmful radiation,

P. Jyothi · S. Koppu (✉)
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India
e-mail: srinukoppu@vit.ac.in

P. Jyothi
e-mail: jyothi.p2020@vitstudent.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_5

use of tobacco, and some hormonal changes in the body. Today it has reached a stage, where artificial intelligence, machine learning, and deep learning techniques are involved for early stage breast cancer diagnosis. Early and timely detection of breast cancer can save 90% of cases. Surgery and radiation are the popular treatment followed for the disease. The disease occurs in the breast, nearby areas, and lymph nodes [2, 3]. Earlier only technique available was X-ray mammography which was used to detect breast cancer. The major drawback of the method is that it cannot be used when the patient has dense breasts moreover the patients are exposed to radiation. These drawbacks are overcome by the methods of magnetic resonance imaging (MRI), ultrasound, and X-ray mammography. It is advisable for people to check for breast cancer if they have a family history of breast cancer or if anyone close to the family is diagnosed with it.

1.1 X-ray Mammography

The most extensively used approach for early breast cancer detection is X-ray mammography. The frequency of the radiation is 30 petahertz to 30 exahertz. An image with a projection of the breast is formed. Spatial resolution is 20 lines pairs/mm. The average detection rate of the method is 78%, and the true positive rate of the method is 98% for the category of 50-plus women with fatty breasts. The efficiency of the method considerably degrades in a younger woman having dense breasts. The existence of radiation can itself be a cause of cancer in this method. The application process is not comfortable for the patients as the tissue has to be pressed hard against a flat surface to get a proper image. In contrast, digital mammography has shown better results in the age category < 50 having dense breasts [4]. Dense parenchyma conceals non-calcified breast cancers under them. In some cases, breast cancers go unnoticed due to the error on part of the physicians. Even with technological advances still, nearly 10% of cancers go unnoticed in the dense regions of the breasts. To increase the sensitivity rate, the detection is also performed with the ultrasound method and MRI.

1.2 Ultrasound Imaging

The sound waves are passed over the breast using a probe that contains transducers. The range of frequency is between 2 and 20 MHz. Ultrasound is preferred over X-ray as the frequency of radiation is much lower. Images are generated on a single pane in real time with the refreshing rate of 25 frames per second. The ultrasound systems can identify breast cancer as small as 3 mm. The detection rate of breast cancer is high when in tandem both mammography and ultrasound are used. The drawback of ultrasound scanning is that it is mostly dependent on the operator. The false positive rate is high due to human error. The drawback of human error is removed in a system with automated screening, but the images have low resolution. The physician performing the ultrasound is the best method, but the learning rate and training required for the physician to master the imaging system require time and a certain level of expertise [5]. The current technology can not only provide a better difference between a cyst and solid areas. The high-frequency transducer generates better shape and margins along with the internal characteristics of solid masses. This leads to a better approximation of benign and malignant cancers. The cost of ultrasound is low, and it provides better guidance for identifying lesions.

1.3 Magnetic Resonance Imaging (MRI)

MRI provides better localization of cancer before surgery. It is a high degree of spatial and temporal resolution, the frequency is in the range of 1-100MHz and the level of radiation is at a safer limit. It detects the tumors in size of mm and is very effective for dense regions.

1.4 Machine Learning Concepts

Machine learning (ML) is making the computer learn by itself using data without explicit programming. The ML algorithms are trained with data and based on the training provided the machine is expected to make predictions.

ML is categorized into two types:

- Supervised learning
- Unsupervised learning.

In supervised learning, the dataset provided to the system contains the class identification field. In unsupervised learning, the class identification field is missing.

Further the supervised learning is divided into

- Classification
- Regression.

The unsupervised learning is further categorized as

- Clustering
- Association.

The ML algorithm is very commonly known as model [5]. There is a model devised for each and every problem to be solved using ML technique. Regression: Regression is used in case of continuous and real variables that affect each other. A small change in one variable has an effect on the other. The regression model predicts a numeric value based on past data. The equation of line $y = \mathrm{MX} + C$ is used by this model. The model tries to minimize the slope and distance of two variables as far as possible. The distance of the plotted data points on the line generated must be the minimum. The predicted line should pass through maximum number of points.

Classification: In classification system, the class field has two or more outcomes like yes, no, cancerous, non-cancerous, etc. During the training phase, the ML system learns all permutations and combinations for each class and predicts in the same manner when it encounters the data it has not learned during training.

Clustering: In clustering, the data that have some sort of similarity are grouped together in single or multiple clusters. K-means is a very popular algorithm in clustering, where *k* identifies the number of clusters required.

Association: The association is rule based to identify which data are associated with one another. It looks for proximity of data based on dataset.

Dimensionality Reduction: In dimensionality reduction, the size or dimensionality of the data is lowered without effecting the quality of the data. The principal component analysis (PCA) is very popular technique to reduce he dimensionality of data. It generates the vectors and looks for spatial similarity between them. The components are generated and the top two components that reveal maximum information on the data is retained.

Ensemble Methods: In ensemble models, different supervised models are combined to provide a higher degree of prediction than the individual model. Here, integration of different models takes place. The popular ensemble model is the random forest model which is a combination of many decision trees. Ensemble models reduce the variation and bias of a single model.

Neural networks and Deep Learning: The neural networks are used to identify the nonlinear patterns in data with the inclusion of parameter layers. The neural network has an input layer, a hidden layer, and an output layer [6, 7]. Figure 1 shows a neural network with a single input layer, a single output layer, and an output layer. In deep learning, the number of layers keeps increasing. The number of hidden layers would be more than one.

Transfer Learning: In transfer learning, the same trained model for some other similar tasks. Once a trained neural network is available, some layers can be passed on to the next layer. This saves both time and computation resources required for training a model.

Reinforced Learning: In reinforced learning, an agent learns from experience. For every task performed by the machine, the agent provides with a reward and for every mistake a penalty. The idea is to maximize the reward after every task and reduce the penalty.

Deep Learning (DL): The ability of DL to handle large amounts of data has made it very popular especially, in the field of pattern recognition. Convolutional neural network (CNN) is the most popular deep neural network. The technique of convolutional is used in CNN. Convolutional is an operation of two functions resulting in a third function that specifies how the modification of shape has taken place [8].

Figure 2 shows a CNN model. The CNN breaks the data into smaller pieces that can be processed easily without losing any features associated with the data. The RGB images have three panes, whereas grayscale images have a single pane [9].

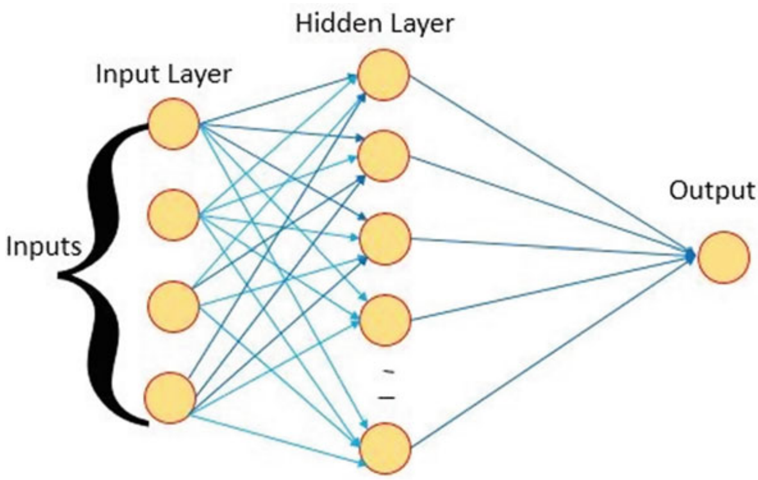


Fig. 1 Neural network

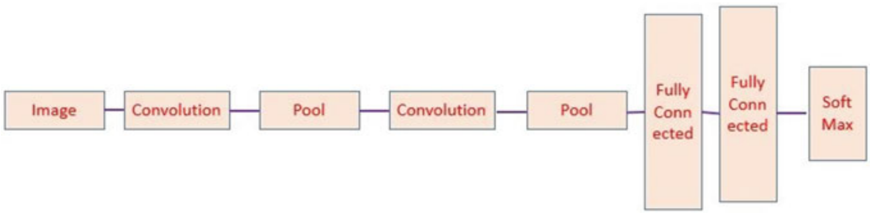


Fig. 2 Convolutional neural network (CNN)

Figure 3 gives the image pattern for RGB and grayscale images. The convolution is performed on the image with the use of a filter or kernel. A 3 × 3 kernel is placed on the image and the convolution takes place. Figure 4 shows an example of the convolution process.

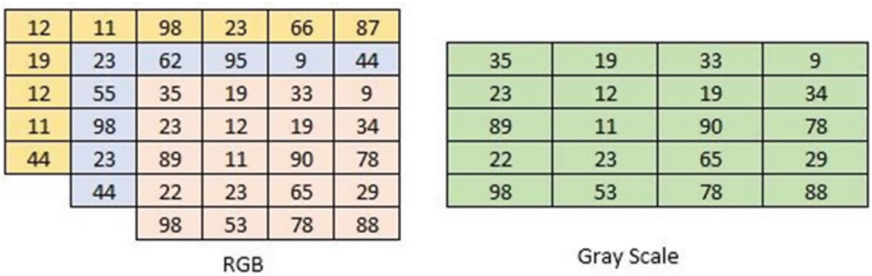


Fig. 3 RGB and grayscale image pattern

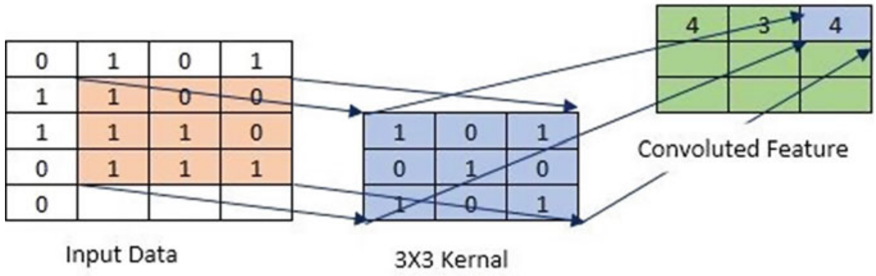


Fig. 4 Convolution process

Pooling Layer: The pooling layer helps to reduce the size of the input image. It is a form of dimensionality reduction. With pooling, the computational requirements also come down. The essential features are maintained but computation is reduced. Max pooling is the maximum pixel value from an area under the image or kernel. Figure 5 shows max pooling and average pooling. The average pooling is the average of the 4 pixels in the area.

Fully Connected Layer: The fully connected layer is found and known as the last layer of CNN and is a feed-forward neural network. The fully connected layer gets the input from pooling or convolutional layers.

Soft Max: Soft Max gives the probabilities of input belonging to a particular class.

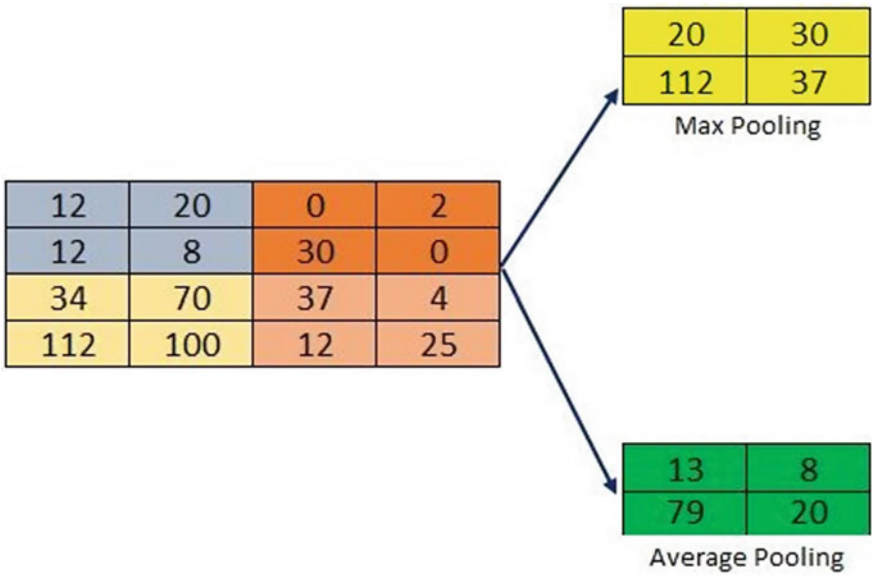


Fig. 5 Max pooling and average pooling

1.5 Literature Review

Reference	Year	Approach	Limitations	Gaps
de Bel et al. [10]	2022	Identify enabled automated quantitative measurement of terminal duct features. Segmentation is performed using CNN	The number of specimens for the study was very less. Only 174 specimens were used by CNN	The system is not fully automated for the prediction process. Uses only standard algorithms like random forest
Jabeen et al. [11]	2022	Breast cancer is detected using fusion of features. The working of the framework is as follows. Convolutional neural network (CNN) data accumulation is executed to enhance the leaning rate of the input model. A trained network is used. The augmented dataset classes are used for modifying the output layer. Transfer learning is implemented, and features are extracted, using two optimizing algorithms, namely reformed differential evaluation (RDE) and reformed gray wolf (RGW) for selecting the main features. The fusion and classification of the features take place	Single dataset of breast images with augmentation is considered	The size of the dataset is small, and the CNN model was not developed from scratch. A pre-trained model was used
Pan et al. [12]	2022	Used prognosis details. The framework consisted of protein expression, somatic mutations, miRNA expression, copy number variation, DNA methylation, and transcription profile. High risk breast cancer patients were easily segregated. The dataset used for the study was TCGA BRCA cohort	Only 501 patient samples from a single dataset were used for the study	Used gene data of breast cancer. No images were used for the study

(continued)

(continued)

Reference	Year	Approach	Limitations	Gaps
Mohamed et al. [13]	2022	Fully system automated breast cancer detection system using U-Net in deep learning. The system structure uses thermography imaging using infrared. First the effected region is extracted segmented using CNN and further processing is performed	The segmentation process is slow, i.e., the time complexity of the algorithm is high. The method has abnormal characteristics when the learning rate increases beyond a value	The DMR-IR dataset is used for the study. 1000 images of total were taken using camera, where 70% of images were taken for training, 15% of the records in the dataset formed the validation records, and the rest 15% were used for testing. The proposed algorithm is not detailed
Wang et al. [14]	2016	Uses metastases from the breast images. The CNN is trained to predict at patch level for the presence of any tumor causing patches. Used threshold-based segmentation on the Camelyon16 dataset. Otsu algorithm implementation takes place after the image is converted to HSV from RGB	The training dataset image were only 90	The basic image processing Otsu algorithm is used. No algorithm is proposed
Akselrod-Ballin et al. [15]	2019	Prediction based on health records and mammography images. Statistical analysis is performed by *t* test and Fisher exact test	Dataset did not consider genetic history among relatives for breast cancer	Other parameters like hormone measurements, breast density and genetic information not considered. Images from only one mammography vendor were included in the study
Dongdong Sun et al. [16]	2018	Prediction system based on integration of multi-dimensional data	Analysis is limited due to small size of dataset	Data belonging to miRNA expression and gene methylation not included in implementation

(continued)

(continued)

Reference	Year	Approach	Limitations	Gaps
Shah et al. [17]	2017	Deep learning model to learn and check the tumor growth and spread at molecular and categorical level. Localization is performed. Biological and data-driven features are analyzed	Algorithm not specified and detailed	Biomarkers related to tumor diagnosis not used in study
Bayramoglu [18]	2016	Classification of histopathology images using CNN. Malignancy is predicted using single stage CNN and image magnification and malignancy is found using multi-stage CNN	Dataset is limited	Deer architectures, normalization, and network splitting not considered

2 Conclusion

Breast cancer is most infected disease among the women in worldwide. The WHO report for the year 2020 claims 685,000 globally lost their life due to breast cancer. It is a disease that has a 90% probability of survival if it can be detected in the early stages. The female gender is prone to breast cancer in the later years of their life, especially after 40. But the chances of getting this disease after attaining puberty are also seen. The previous methods for the detection of breast cancer such as X-ray mammography, ultrasound imaging, and magnetic resource imaging have their advantages and drawbacks with all the technologies available. Deep learning using complex machine learning classifiers has made a remarkable impact, and with the use of the latest technologies and techniques in deep learning, the detection of breast cancer was made easily detected with images.

References

1. Islam MS, Kaabouch N, Hu WC (2013) A survey of medical imaging techniques used for breast cancer detection. In: IEEE international conference on electro-information technology. EIT, Rapid City, SD, USA, pp 1–5. https://doi.org/10.1109/EIT.2013.6632694
2. Singh AK (2015) A novel approach for breast cancer detection and segmentation in a mammogram. Procedia Comput Sci 7
3. Ehteshami Bejnordi B et al (2017) Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 318(22):2199. https://doi.org/10.1001/jama.2017.14585
4. Nagi J, Abdul Kareem S, Nagi F, Khaleel Ahmed S (2010) Automated breast profile segmentation for ROI detection using digital mammograms. In: IEEE EMBS conference on biomedical

engineering and sciences (IECBES), Kuala Lumpur, Malaysia, pp 87–92. https://doi.org/10.1109/IECBES.2010.5742205
5. Prasad SN, Houserkova D (2007) The role of various modalities in breast imaging. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub 151(2):209–218. https://doi.org/10.5507/bp.2007.036
6. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Networks 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003
7. Mesnil G et al (2015) Using recurrent neural networks for slot filling in spoken language understanding. IEEE/ACM Trans Audio Speech Lang Process 23(3):530–539. https://doi.org/10.1109/TASLP.2014.2383614
8. Langkvist M, Karlsson L, Loutfi A (2014) A review of unsupervised feature learning and deep learning for time-series modeling. Pattern Recogn Lett 42:11–24. https://doi.org/10.1016/j.patrec.2014.01.008
9. Pratiwi M, Alexander, Harefa J, Nanda S (2015) Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network. Procedia Comput Sci 59:83–91. https://doi.org/10.1016/j.procs.2015.07.340
10. de Bel T et al (2022) Automated quantification of levels of breast terminal duct lobular (TDLU) involution using deep learning. npj Breast Cancer 8(1):13. https://doi.org/10.1038/s41523-021-00378-7
11. Jabeen K et al (2022) Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. Sensors 22(3):807. https://doi.org/10.3390/s22030807
12. Pan X, Burgman B, Sahni N, Yi SS (2022) Deep learning based on multiomics integration identifies potential therapeutic targets in breast cancer. Bioinformatics. https://doi.org/10.1101/2022.01.18.476842
13. Mohamed EA, Rashed EA, Gaber T, Karam O (2022) Deep learning model for fully automated breast cancer detection system from thermograms. PLoS ONE 17(1):e0262349. https://doi.org/10.1371/journal.pone.0262349
14. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH (2016) Deep learning for identifying metastatic breast cancer. [cs, q-bio]. [Online]. Available: http://arxiv.org/abs/1606.05718. Accessed: 09 Feb 2022
15. Akselrod-Ballin A et al (2019) Predicting breast cancer by applying deep learning to linked health records and mammograms. Radiology 292(2):331–342. https://doi.org/10.1148/radiol.2019182622
16. Sun D, Wang M, Li A (2019) A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data. IEEE/ACM Trans Comput Biol Bioinf 16(3):841–850. https://doi.org/10.1109/TCBB.2018.2806438
17. Shah M, Rubadue C, Suster D, Wang D (2016) Deep learning assessment of tumor proliferation in breast cancer histological images. [Online]. Available: http://arxiv.org/abs/1610.03467. Accessed: 09 Feb 2022
18. Bayramoglu N, Kannala J, Heikkila J (2016) Deep learning for magnification independent breast cancer histopathology image classification. In: 23rd international conference on pattern recognition (ICPR), Cancun, pp. 2440–2445. https://doi.org/10.1109/ICPR.2016.7900002

Ensemble Decision Fusion of Deep Learning Classifiers for Heart Disease Classification

K. Shilpa and T. Adilakshmi

Abstract Earlier detection of heart disease has gained wide research interest due to its higher mortality and challenges its accurate prediction. Computer-aided diagnosis is of great assistance to physicians in making accurate decision on heart diseases. Machine learning algorithms are the core of computer-aided diagnosis. Various machine learning algorithms both traditional and deep learning have been proposed for heart disease diagnosis. Recently, deep learning algorithms have been increasingly used in computer-aided diagnosis due to ability to learn intricate features without need for handcrafted feature selection. This work proposes an ensemble of deep learning classifiers for classification of heart diseases from ECG data.

Keywords Machine learning · Deep learning · Heart diseases prediction · CNN and ECG data

1 Introduction

Cardiovascular diseases claim more than 17 million liver every year worldwide. Almost 80% of them are due to heart attack and stroke. India has fastest growing number of cardiovascular disease patients in recent years due to drastic shift in quality of life and food habits. India has more than 30 million heart disease patients and over two lakh open heart surgeries are performed each year. A more worrying concern is that number of patients requiring coronary interventions is rising at alarming rate of 20–30% every year. Many of heart related deaths could have been avoided if it

K. Shilpa (✉)
Department of CSE, UCE, OU, Hyderabad, India
e-mail: shilpak@osmania.ac.in

Department of CSE, CMR Technical Campus, Hyderabad, India

T. Adilakshmi
Department of CSE, Vasavi College of Engineering, OU, Hyderabad, India
e-mail: t_adilakshmi@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_6

has been diagnosed in time and effective treatment procedures have been started in advance. Heart disease diagnosis is complicated and depends on doctor expertise. Due to diagnosis cost and time involved, patients hardly go through regular diagnosis procedure. This necessitates use of low-cost automated diagnosis system which can save cost and time for the patients. Use of data mining is explored in this work for providing a low-cost automated diagnosis system for patients. Huge amount of data is collected from patients by healthcare institutions nowadays. Mining on these data, various inferences on disease prediction and categorization can be done. The problem in mining on these huge data is selection of attributes relevant for classification of the disease. Once the features are selected, the classification model is built using various algorithms like KNN, neural network, Naïve Bayes, SVM, etc.

Recently, deep learning is proposed as an effective means to replace hand crafted features. Deep learning models are able to learn more intricate features from the data, and these features have high disease discriminating ability. This work proposes a decision-level fusion of classification based on deep learning features to increase the accuracy of heart disease classification. ECG signals are preprocessed with various wavelet filters and passed as input to trained convolutional neural network to classify heart diseases. The results are then ensembled to provide the final classification result. Through ensembling of results from multiple deep learning models, the accuracy of heart disease classification increases.

Paper organization is as follows. Section 2 presents the existing solutions for heart disease classification. Section 3 presents the proposed deep learning ensemble decision fusion. Section 4 presents the results of proposed solution and comparison to existing works. Section 5 presents the conclusion and scope for future research.

2 Literature Survey

Sultana et al. [1] analyzed the performance of five classifiers. The classifier were Bayes Net, SMa, KStar, MLP and J48. From their analysis, Bayes Net and SMa classifiers are the optimum among the all five classifiers. Their analysis is based on dataset collected at particular time and not on continuous time series data.

Gandhi [2] detailed different classification methods for heart disease prediction. They studied decision trees, neural networks and Naïve Bayes classifier for prediction. The dataset attribute selection was not given attention in this work.

Colak et al. [3] studied the use of knowledge discovery process to analyze stroke using ANN and SVM classifiers. Feature selection was done using Cramer's V test to select the attributes. Around 80% accuracy was achieved in this method. One of important inferences in their analysis is feature selection gives better accuracy.

Acharya et al. [4] proposed a heart disease classification system. Features are extracted from ECG signals and used for classification. Spectral features are extracted and the dimensions are reduced using principal component analysis. Then, coefficients are ranked using multiple ranking methods. KNN and decision tree models are built using ranked features to get the highest accuracy. The classification cannot predict initial stages of disease.

Hashi et al. [5] implemented an expert system using decision tree and KNN classifier to predict disease. They used Pima Indians Diabetes Database to predict the diabetes. The disadvantage in this work is that it is based on snapshot data and not on continuous time series data.

Shah et al. [6] proposed a decision support systems for disease classification with more focus on feature selection. Probabilistic principal component analysis (PPCA) was used for feature extraction. It extracted features with high covariance and feature dimension is also reduced in this work. On the dimension reduced data, RBF neural network is built to predict the disease. With usage of feature selection accuracy improved, but the dataset is not continuous and it is only snapshot.

Saqlain et al. [7] proposed a risk model for heart disease. This model can predict 1-year or more survival for heart failure diagnosed patients. From the unstructured medical records, certain attributes are extracted using search, and then, multinominal Naïve Bayes (NB) is used for classification. The approach does not address missing value problem. Feature selection is not given attention.

Dwivedi [8] analyzed heart disease prediction problem using six machine learning techniques. Thirteen distinct attributes from dataset are taken for analysis. StatLog heart disease dataset from UCI machine learning repository is taken and classifier models built on those dataset to predict heart disease. But the work does not give attention to feature set selection and continuous data.

Buchan et al. [9] proposed a prediction system for coronary heart disease. It used narrative medical histories and extracted features from it for training and classification. From the unstructured text input, features are extracted using natural language processing techniques, and from it PCA, mutual information filter employed to reduce feature dimension size. SVM classifier is then built to classify. The system accuracy depends on narration ability of patients, and when missing information in narration, the accuracy is very low.

Alickovic and Subasi [10] extracted DWT features form ECG signals and trained a decision tree classifier to classify arrhythmia. The accuracy of the method is not satisfactory as no feature dimensionality reduction procedures are followed. Also data is snap shot based and not on continuous based.

Jabbar et al. [11] proposed a combined classifier to classify heart diseases. It used KNN and genetic algorithm as a combination. Genetic algorithm removed redundant and irrelevant attributes, and for ranking the attributes which contribute more toward classification. Low ranked attributes are removed and KNN model is built on it for classification. By using genetic search classification, accuracy is improved in this approach.

Kumar and Inbarani [12] devised a heart disease classification using ECG signals. Signal denoising is done using wavelet filters, and on the filtered signal, discrete wavelet transform (DWT)-based features are extracted. The method is only for data collected at a particular snapshot.

Mustaqeem et al. [13] proposed a method based on selecting best features with random forest algorithm. After selected features are extracted, different machine learning techniques are applied to predict heart disease. They were able to achieve 75% accuracy using this method.

Khanna et al. [14] proposed a machine learning classification for heart diseases. They used logistic regression, support vector machines and neural networks for classification. Cleveland heart disease dataset was used for analysis. But the accuracy is not good enough in this approach and it is not on continuous data.

3 Ensemble Decision Fusion

The ensemble decision fusion architecture is given in Fig. 2. The system does an ensemble classification on four different deep learning classifiers. Four different deep learning classifiers used in this work are (Table 1).

Each of the classifiers is detailed below.

3.1 Deep Learning ECG-Based CNN (C1)

The deep learning ECG-based CNN takes the raw ECG as input, preprocess it and classifies it to two classes of heart disease or no heart disease. ECG signals are segmented in epochs of 5 min of 3000 samples for duration of 90 min in a non-overlapping mode. Hence, there are totally 16 segments in each training sample. ECG segments are preprocessed before learning features from it using CNN. Each ECG segment of 5 min duration is individually mean centered and normalized to a standard deviation of 1. From the ECG segments, power line interference at 50 Hz is removed using notch filter and then mean of the signal is removed. Convolutional neural network (CNN) is a deep learning model popular for its ability to learn suitable spatially invariant features. The segments of ECG are assembled as 2D (3000 * 16) matrix with each row being a segment. The 2D matrix is passed as input to the CNN instead of manually extracted features, so that CNN can learn the patterns for detection of seizures. A univariate feature extraction topology is designed for the CNN. The convolution, nonlinear activation and pooling layer in CNN extract features on different time scales. Following are the functions used at different layers in the CNN (Table 2).

The design of the CNN for the case of C1 is given in Fig. 1 (Fig. 2).

Table 1 Deep learning models

C1	Deep learning ECG-based CNN
C2	DFT feature-based CNN classifier
C3	Tunable-Q wavelet transform (TQWT) feature-based CNN classifier
C4	Db4-based CNN classifier

Table 2 Activation functions in layers

Layer	Function
Convolutional layer	Nonlinear activation
Pooling layer	Rectified linear units
Output layer	Sigmoid function

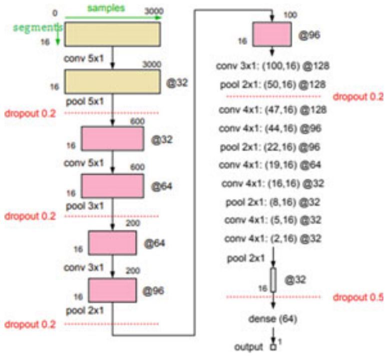


Fig. 1 CNN model for classification

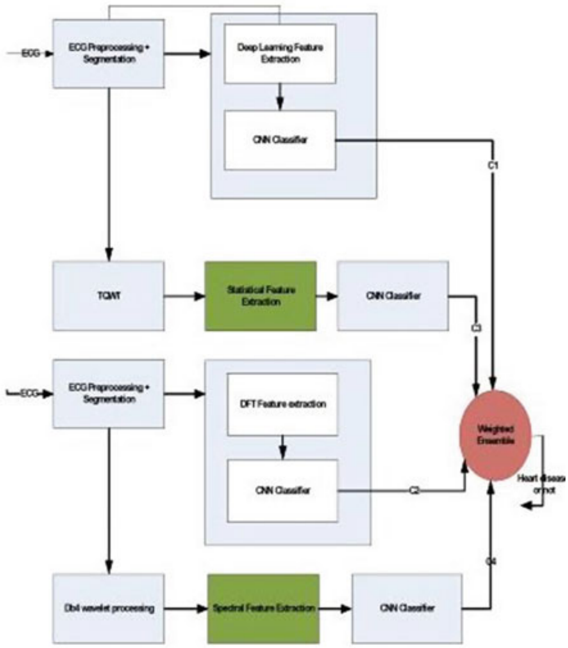


Fig. 2 Ensemble decision fusion

The training sample of 16 rows each with 3000 columns is processed with 32 layer neural network. The receptive field of deep neurons grows on the time scale due to subsequent convolution and pooling layers. The features got as out from the last convolution and pooling layer is then passed to fully connected layer for classification of heart disease.

3.2 DFT Feature-Based CNN Classifier (C2)

DFT transforms a periodic signal in terms of sequence of discrete Fourier transform coefficients. By applying the Fast Fourier transformation, the coefficients of time series data over frequency domain are captured for every 5 min time window. On the coefficients resulting from Fast Fourier transformation, following statistical features are extracted (i) maximum value (ii) minimum value, (iii) mean and (iv) standard deviation. A matrix of statistical features for time window is constructed and passed to CNN to classify heart disease.

3.3 TQWT Feature-Based CNN Classifier (C3)

TQWT is a wavelet transform and it suits for extracting features from oscillatory signals like ECG. TQWT uses multiple filter banks in sequence ordered in form of low pass to high pass.

TQWT is fully discrete wavelet transform. Each filter bank has a low pass filter $H_0^j(w)$ and high pass filter $H_1^j(w)$. The filters are defined as

$$H_0^j(w) = \begin{cases} \prod_{m=0}^{j-1} H_0\left(\frac{w}{\alpha^m}\right), & |w| \le \alpha^j \pi \\ 0, & \alpha^j \pi < |w| \le \pi \end{cases}$$

$$H_1^j(w) = \begin{cases} H_1\left(\frac{w}{\alpha^{j-1}}\right) \prod_{m=0}^{j-2} H_0\left(\frac{w}{\alpha^m}\right), & (1-\beta)\alpha^{j-1}\pi \le |w| \le \alpha^{j-1}\pi \\ 0, & \text{for other } w \in [-\pi, \pi] \end{cases} \text{ for}$$

where

$$H_0(w) = \theta\left(\frac{w + (\beta - 1)\pi}{\alpha + \beta - 1}\right)$$

$$H_1(w) = \theta\left(\frac{\alpha\pi - w}{\alpha + \beta - 1}\right)$$

$\theta(w)$ is the Daubechies filter frequency response.

J is the number of decomposition levels and it is set as 3 for this work.
α is the low pass scaling factor and it is set as 0.6 for this work.
β is the high pass scaling factor and it is set as 1 for this work.

The ECG signal is segmented in epochs of 5 min of 3000 samples for duration of 90 min in a non-overlapping mode. TQWT transform is done on each of the segmented and from the wavelet coefficients following statistical features are extracted

- Mean
- Standard deviation
- Variance
- Kurtosis
- Percentile (per 75).

After the features are extracted for each of the segments, it is arranged as 2D matrix (16 * 5, 16 rows one for each segment and 5 columns one column for each feature). A convolutional neural network (C3) with one convolution and one pooling layer is trained with these features to classify the heart disease based on ECG features.

3.4 *Db4-Based CNN Classifier (C4)*

ECG signals are segmented in epochs of 5 min of 3000 samples for duration of 90 min in a non-overlapping mode. A variant of Daubechies wavelet called as db4 wavelets is applied on each of the segments. The choice of db4 is due to its peculiarity of maximum number of vanishing moments and its powerful scaling. Following features are extracted from the wavelet coefficients.

1. Maximum power value from power spectral analysis
2. Mean
3. Kurtosis
4. Skewness
5. Total harmonic distortion.

After the features are extracted for each of the segments, it is arranged as 2D matrix (16 * 5, 16 rows one for each segment and 5 columns one column for each feature). A convolutional neural network (C4) with one convolution and one pooling layer is trained with these features to classify the heart disease based on ECG features.

3.5 *Ensemble Decision Fusion*

From the results of C1, C2, C3 and C4 classifier, ensemble decision fusion is done to get the final classification result. Weighted ensemble of results from the four classifiers (C1, C2, C3 and C4) is done to detect heart disease. Each deep learning classifier provides a binary decision of whether heart disease is detected or not. The decision of each of classifier is weighted ensembled to arrive at final result. The weight is calculated as

$$w(C_i) = \log \frac{1 - \text{error}(C_i)}{\text{error}(C_i)}, \quad 1 \leq C_i \leq 4$$

Trail run is done to detector the error rate for each of the classifier. The results of C1, C2, C3 and C4 are used in ensemble model. The behavior of the weighted sampling is illustrated with an example. Say for the error rate given in Table 3, the calculated weights are given in Table 3.

For a sample classifier result in Table 4, the decision fusion is done as below.

The weights of fusion for heart disease class are 0.78 and the weights of fusion for non-heart disease class are 1.3. The weights for non-heart disease class (1.3) are higher than weights of heart disease class (0.78), so the decision is made as non-heart disease for this case.

Following are the novel contributions of this work

(i) Ensembling of multiple deep learning models to improve the heart disease recognition accuracy
(ii) Wavelet processing based convolutional neural network model for long term temporal analysis of ECG patterns.

Table 3 Example error rate

Classifier	Error rate	Weights
C1	0.25	0.47
C2	0.14	0.78
C3	0.30	0.36
C4	0.25	0.47

Table 4 Ensemble decision fusion example

Classifier	Heart disease (*a*)	No heat disease (*b*)	Weighted score for heart disease (*a* * weights)	Weighted score for non-heart disease (*b* * weights)
C1	0	1	0	0.47
C2	1	0	0.78	0
C3	0	1	0	0.36
C4	0	1	0	0.47
Total weights			0.78	1.3

4 Results

The proposed solution was tested on South Africa Heart Disease Dataset [15]. The dataset consists of heart disease samples in high risk populations. The dataset is collected from patients who have undergone chronic health disease tests, blood pressure reduction treatment and other programs devised for reduction of health risks.

The performance of the proposed solution is tested in terms of standard metrics of (i) accuracy, (ii) sensitivity, (iii) specificity (iv) false detection ratio (FDR). The performance of proposed solution is compared against SVM-based classification [15] and inception residual network proposed by Ni et al. [16]. The performance was measured with fivefold cross-validation with ratio of 80:20 for training and testing.

The highest value for sensitivity, specificity and FDR of 97.8%, 97.4% and 1 is achieved in proposed ensemble decision fusion compared to [16, 17] as in Table 5.

The ROC curve obtained for all 3 solutions is given Fig. 3.

Maximum AUC of 0.97 is obtained in the proposed multi-modal ensemble fusion and its highest compared to [16, 17]. The comparison results for accuracy are given in Fig. 4.

Highest accuracy of 98.4% is obtained in the proposed multi-modal fusion compared to [16, 17]. The comparison of accuracy obtained in each of classifier

Table 5 Performance comparison

Method	Sensitivity	Specificity	FDR
Papadogiorgaki et al. [17]	86.9	94.4	3
Ni et al. [16]	86.5	80.0	3
Proposed	97.8	97.4	1

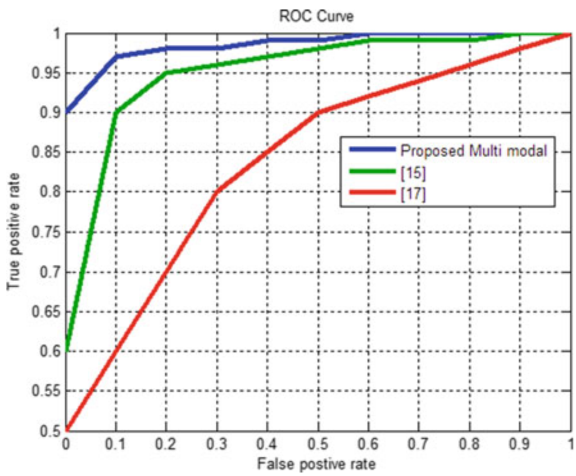


Fig. 3 Comparison of ROC

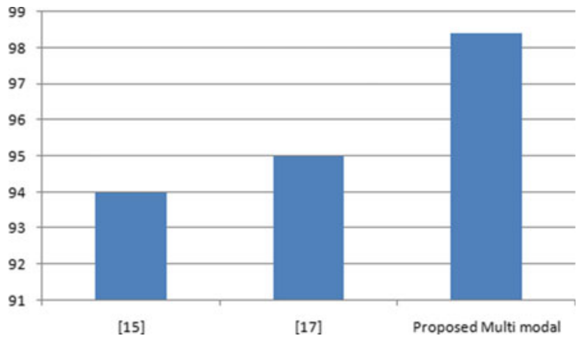


Fig. 4 Comparison of accuracy

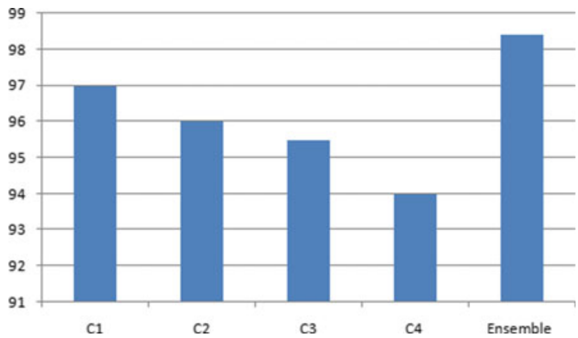


Fig. 5 Classifier wise accuracy

(C1, C2, C3 and C4) and the ensemble in the proposed multi-modal ensemble fusion is given in Fig. 5.

From these results, it can be seen that accuracy in C1 and C2 is higher C3 and C4. Also the accuracy due to ensemble is higher than that of individual classifiers.

5 Conclusion

This work proposed an ensemble decision fusion of deep learning features for heart disease classification. ECG signals are processed with different wavelets, and features extracted are passed to more intricate feature-based classification by convolutional neural networks. The results of the each of the convolutional neural networks are ensembled to provide higher accuracy of heart disease classification. The proposed solution is able to achieve about 98% accuracy and it is at least 3% higher compared to earlier works.

References

1. Sultana M, Haider A, Uddin MS (2016) Analysis of data mining techniques for heart disease prediction. In: 3rd international conference on electrical engineering and information communication technology, iCEEiCT 2016, 2017
2. Gandhi M (2015) Predictions in heart disease using techniques of data mining. In: International conference on futuristic trends on computational analysis and knowledge management
3. Colak C, Karaman E, Turtay MG (2015) Application of knowledge discovery process on the prediction of stroke. Comput Methods Programs Biomed 119(3):181–185
4. Acharya UR et al (2017) Application of higher-order spectra for the characterization of Coronary artery disease using electro-cardiogram signals. Biomed Signal Process Control 31:31–43
5. Hashi EK, Zaman MSU, Hasan MR (2017) An expert clinical decision support system to predict disease using classification techniques. In: International conference on electrical, computer and communication engineering. pp 396–400
6. Shah SMS, Batool S, Khan I, Ashraf MU, Abbas SH, Hussain SA (2017) Feature extraction through parallel probabilistic principal component analysis for heart disease diagnosis. Phys A Stat Mech Appl 482:796–807
7. Saqlain M, Hussain W, Saqib NA, Khan MA (2016) Identification of heart failure by using unstructured data of cardiac patients. In: 45th international conference on parallel processing workshops. pp 426–431
8. Dwivedi K (2016) Performance evaluation of different machine learning techniques for prediction of heart disease. Neural Comput Appl 1–9
9. Buchan K, Filannino M, Uzuner Ö (2017) Automatic prediction of coronary artery disease from clinical narratives. J Biomed Inform 72:23–32
10. Alickovic E, Subasi A (2016) Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier. Patient Facing Syst J Med Syst 40:108–120
11. Jabbar M, Deekshatulua B, Chandra P (2013) Classification of heart disease using k-nearest neighbor and genetic algorithm. Procedia Tech 10:85–94
12. Kumar SU, Inbarani HH (2017) Neighborhood rough set based ECG signal classification for diagnosis of cardiac diseases. Soft Comput 21(16):4721–4733
13. Mustaqeem A, Anwar SM, Majid M, Khan AR (2017) Wrap-per method for feature selection to classify cardiac arrhythmia. In: Proceedings of the 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC). pp 3656–3659
14. Khanna D, Sahu R, Baths V, Deshpande B (2015) Comparative study of classification techniques (SVM, logistic regression and neural networks) to predict the prevalence of heart disease. Int J Mach Learn Comput 5(5):414–419
15. South Africa Heart Disease Dataset Source. https://web.stanford.edu/~hastie/ElemStatLearn//data.html
16. Ni J et al (2021) Multi-class cardiovascular disease detection and classification from 12-lead ECG signals using an inception residual network. In: 2021 IEEE 45th annual computers, software, and applications conference (COMPSAC). pp 1532–1537
17. Papadogiorgaki M et al (2021)Heart rate classification using ECG signal processing and machine learning methods. In: IEEE 21st international conference on bioinformatics and bioengi-neering (BIBE). pp 1–6

Implementation of Automatic Vending Machine Using FPGA

P. Anuradha, K. Rajkumar, Ch. Navitha, and M. Jithender Reddy

Abstract Vending machine is used to dispense the items from the machine for given amount. This machine is implemented using FPGA board. FPGA board-based machine gives us fast response and it is reprogrammable. This machine takes money such as coins as input and gives out the product for which the money is inserted. If the amount inserted is greater than cost of product, then the vending machine gives out the remaining change. This machine also allows a cancel button to cancel the product if we don't need it. The implementation of paper is using Verilog HDL in Xilinx ISE Simulator tool and using Xilinx Spartan-3 FPGA development board.

Keywords FPGA · Verilog · HDL · Xilinx spartan-3

1 Introduction

Vending machines are designed to provide different products such as grocery, jewelers, medicines and drinks. The first vending machine was introduced in London. It is used for selling postcards. These are implemented using different technologies such as microcontrollers and CMOS but we are introducing a vending machine that works using FPGA which is more flexible, reprogrammable and less power consumption. In other technologies, if anyone wants to change the design, they need to change total hardware structure, but in FPGA, we can reprogram it. We can use this machine in many places like offices, hospitals, marts, etc. [1].

P. Anuradha (✉) · Ch. Navitha
Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: anu.rnitw99@gmail.com

K. Rajkumar
SR University, Warangal, Telangana, India

M. Jithender Reddy
Vasavi College of Engineering, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_7

The machines normally work, when some amount is put in a slot [2]. Then, a button is to be pushed. If there is sufficient amount in the machine, then the product which is selected by the purchaser will be dropped into a tray. This machine is seen by an administrator so no more workers are needed. It can work 24/7 whenever in occasions, holidays and every day.

FPGA has many applications such as computer hardware emulsion, prototyping, cryptography, voice recognition, filtering and communication encoding and medical imaging.

Block diagram is shown in Fig. 1 which has coin sensors, Relays, FPGA Xilinx spartan-3 board and Power supply. Spartan-3 development board is designed to prototype most common FPGA application [3]. This FPGA board is used to implement hardware digital design to advanced controllers. In FPGA, support circuits are included so that countless designs can be created without using any other external components. Relay is a switch that is electrically operated and used for the inputs for the machine. It contains set of input terminals and a set of operating contact terminals. LCD display is used to display the information in the output [4]. Coin Sensor is used to sense coins that identify the coins by their thickness, diameter and fall time of the coins. It is fully programmable so you are not limited to any particular currency [5].

The main theme of this paper is providing a vending machine that provides different products to the people using simple steps. Here, we are providing 4 items in the machine with different price for each product. This machine also provides change to the customers if they insert extra money for the product. We provide an extra option named cancel which is used to cancel the product if the customer doesn't

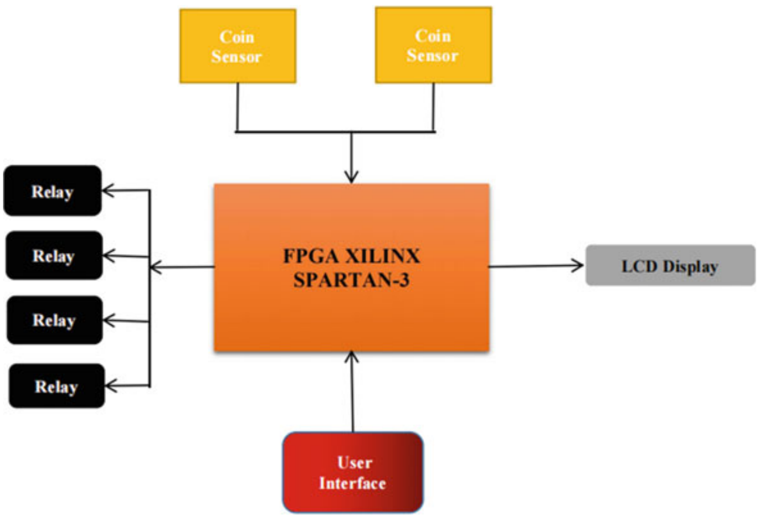


Fig. 1 Block diagram of vending machine using FPGA

need the product after the selection. They can also withdraw money using this cancel button [6].

2 Proposed Work

A vending machine is used to dispense different types of products after an amount is deposited into the machine [7]. After doing many researches, we came to know that the main problem with other machines is having complex hardware structure, high power consumption and time consuming. To overcome this, we are implementing the paper using FPGA board. These machines first introduced in government transport and hospitality to provide all the convenience to the customers by providing essential products. It is a vending machine that provides the products when an amount is inserted into the machine.

After money is inserted into machine, there is some mechanical process to get the product. The machine is releasing out the product so it falls into an open compartment like structure fixed to that machine at the bottom. This takes place after a releasing or opening of a knob. The first vending machine came into existence in first century from Roman Egypt which is used to dispense holy water so that people were not allowed to use more holy water. These machines are implemented using many technologies like microcontrollers, CMOS, SED, RFID, Arduino, HTML, Java, etc. These machines implemented using singled electron encoded logic (SEEL) are much time consuming [8].

CMOS, SED, microcontroller-based vending machines are also available with best security and power consumption but the main disadvantage is the system works very slow and less efficient. Some machines use RFID technology which is least cost, easy to use but the main disadvantage of this technology is it only detects RFIG tags [9]. Arduino-based machine also introduced but the disadvantage is power consumption and security.

Embedded system-based vending machines are also designed to improve the efficiency of the machines [10]. Mainly these machines are introduced to lower the workload for humans in marts, coffee shops, medical shops, etc. Considering a coffee shop, it is very difficult to workers to make coffee for each person and make billing at a time. Vending machines are vending machines that are designed to dispense different products using different technologies. But the main aim of our paper is implementing a vending machine using FPGA Xilinx spartan-3 board.

FPGA contains semiconductor devices and consists of configurable logic blocks (CLBs) and connected via programmable interconnects [11]. The inputs and outputs of FPGA are ones and zeros. FPGA is used for papers, where the hardware connection can be changed according to the change in the paper. FPGA programming consists in creating hardware architecture that will execute a requested algorithm and describe it in a hardware description language (HDL).

Xilinx ISE simulator tool enables the developer to synthesize their designs, time analysis, RTL diagrams and configure target design to the programmer [12]. Xilinx's contains patented algorithms to synthesis allow designs to run 30% faster than competing programs and allows greater logic density which reduces paper time and costs. This simulator follows some steps in simulating.

Firstly, we need to do logical verification, to make sure that the module expresses the expected results.

Second, behavioral verification to verify logical and timing issues. Third, post place and routing simulation to verify behavior within the reconfigurable logic of FPGA. The machines that are developed by FPGA are flexible, programmable and fast response. If the designer wanted to change any or add any feature, then no need to change whole hardware structure in the circuit because FPGA is reprogrammable so we can program it and add feature.

3 Design Methodology

Design process of vending machine using FPGA is shown in Fig. 2. The machine we implemented is dispensing four different products. Each product is considered as one state so there are four states. When reset is equal to zero, then the machine is said to be in initial state. When reset equal to 1, the user can select the product that they need [13]. Let us assume the four products such as product 1, product 2, product 3 and product 4. First, the availability of product is checked by the machine when user selects the product. If the product is not available, then machine goes to its initial state.

If the product is available, then the user intimated to insert the amount for the product. If in any circumstances, the users do not need that selected product then the

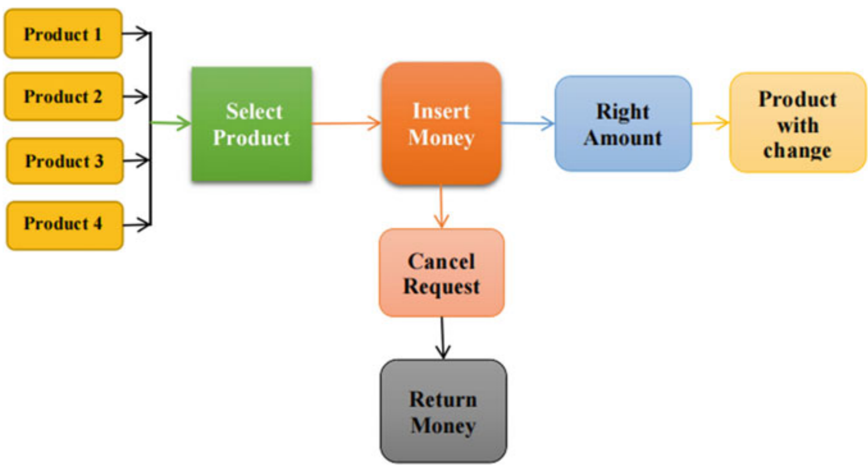


Fig. 2 Design process of vending machine using FPGA

user can click a cancel button so that the money or amount inserted for the product is returned to the user [14].

Otherwise, the product is dispensed from the machine, and if the amount inserted is not sufficient, it will not dispense the product and goes to initial state. If the amount inserted is greater than the cost of the product, then the machine gives out the remaining change. And again it goes to initial state. This process repeats for every time user selects the product.

A cancel is used when the customers do not need the product that they selected, if the user has inserted the amount and want to change the product or else not needed any product, the amount they give into the machine is returned back automatically to the customer. Please note that, if your email address is given in your paper, it will also be included in the metadata of the online version.

4 Results and Discussions

The simulation is done separately and at last combined all the results. The products one by one selected according to the requirements, and the output is verified. Simulation of the product is shown in Fig. 3.

In code simulation, first we have selected a product named string. Reset always becomes 1 when the machine starts its procedure. If reset is zero, it represents the machine is in its initial state. When coffee is selected, the output wave of the coffee becomes high as it represents that it is selected.

We have given different binary values to separate products and the amount of Rs. 10, Rs. 20, etc. In Fig. 4, we can see the amount or money binary value is 10.

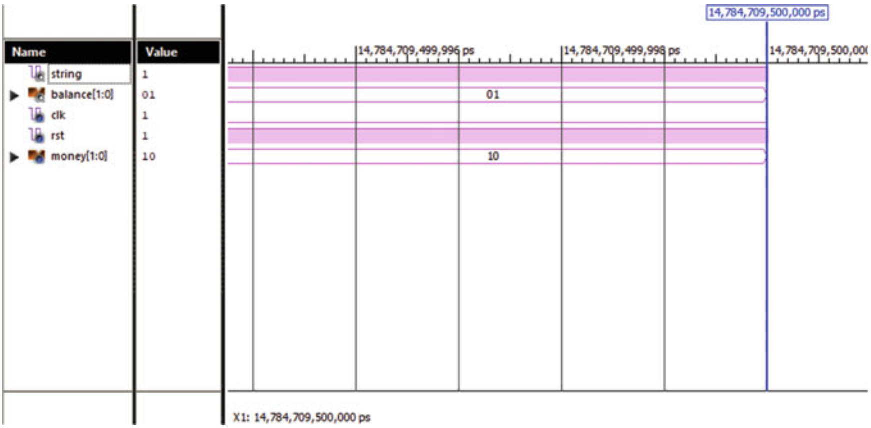


Fig. 3 Simulation waveform of string selection

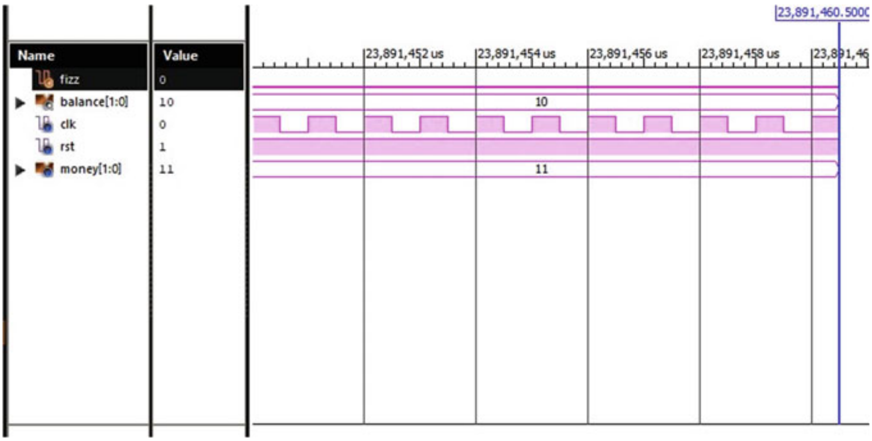


Fig. 4 Simulation waveform of fizz selection

Figure 4 shows the simulation results of selection of another product named fizz. The process is same as the selection of string but the assigned binary value is different. The money binary value is changed to 11.

Figures 5 and 6 show the simulation results of selection of products soda and diet coke. The process is same as selection of other products, and there is only change in the states and money returned.

Figure 7 shows the results of all the products and remaining amount, balance and change. All the simulation results are verified and shown in the figures. In similar fashion, all the products are delivered according to the amount and selection of the

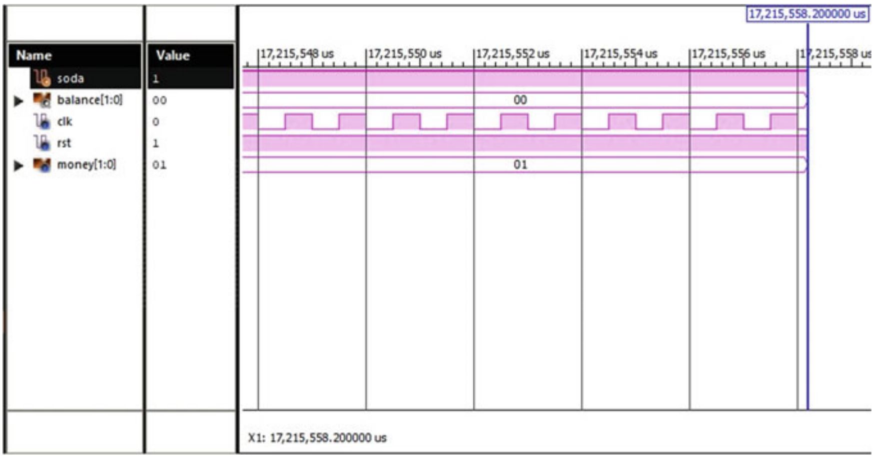


Fig. 5 Simulation waveform of soda selection

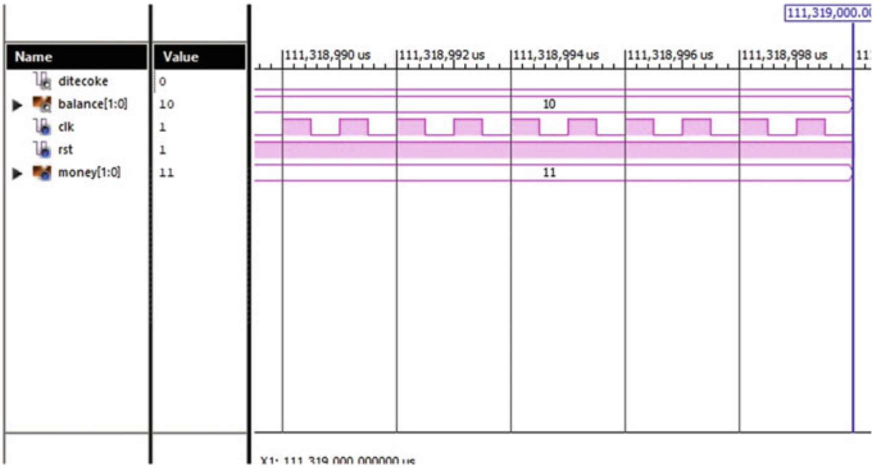


Fig. 6 Simulation waveform of diet coke selection

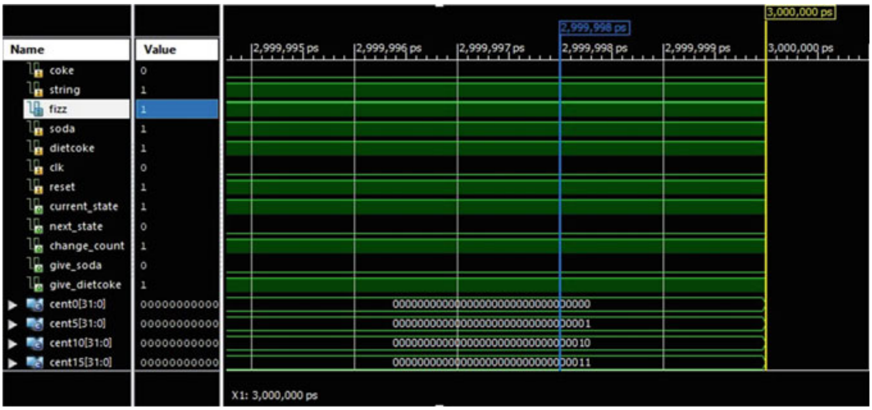


Fig. 7 Simulation waveform of all the products

product. After successful implementation and product delivery, the user provides cancel button.

5 Conclusion

This proposed paper of implementation of vending machine using FPGA is designed using Verilog programming and simulation is done by Xilinx software. It is observed that in many cases, FPGA is more flexible, faster response and low power consumption. The proposed FPGA-based vending machine is used in many applications and users can add the products or change the products based on the requirements. FPGA normally has many applications such as cryptography, prototyping, filtering and voice recognition. Our results clearly show the efficiency of FPGA and easy use of the vending machine for the user.

References

1. Arabelli RR, Rajababu D (2019) Transformer optimal protection using internet of things. Int J Innovative Technol Exploring Eng 8(11):2169–2172
2. Arabelli RR, Revuri K (2019) Fingerprint and Raspberri Pi based vehicle authentication and secured tracking system. Int J Innovative Technol Exploring Eng 8(5):1051–1054
3. Anuradha P (2019) An intelligent computing framework for multicore heterogeneous embedded architectures. IJAST 28(17):701–708
4. Anuradha P, Rallapalli H, Narasimha G, Ahmed SM (2015) Efficient workload characterization technique for heterogeneous processors. In: IEEE international advance computing conference (IACC). IEEE, pp 812–817
5. Anuradha P, Rallapalli H, Narsimha G (2018) Energy efficient scheduling algorithm for the multicore heterogeneous embedded architectures. Des Autom Embed Syst 22(1):1–2
6. Anuradha P, Rallapalli H, Narsimha G (2018) Versatile intelligent ELM algorithm for workload characterization. J Adv Res Dyn Control Syst 10:177–184
7. Anuradha P (2019) Software and hardware tool for the development of embedded software and a study on applications and characteristics of embedded system. Int J Adv Sci Technol 28(17):1–8
8. Arabelli R, Anuradha P, Rajkumar K, Yedulapuram S (2020) Smart street light management system for conservation of electrical energy. In: IOP conference series: materials science and engineering, vol 981, no 3. IOP Publishing, p 032041
9. Anuradha P, Arabelli R, Rajkumar K, Ravichander J (2020) Microcontroller based bi-directional vehicle counter and automatic gate controlling system. In: IOP conference series: materials science and engineering, vol 981, no 3. IOP Publishing, p 032023
10. Anuradha P, Arabelli RR, Rajkumar K, Ravichander J (2020) Microcontroller based monitoring and controlling of LPG leaks using Internet of Things. In: IOP conference series: materials science and engineering, vol 981, no 3
11. Samdani J, Anuradha P, Rao H (2018) VLSI architecture of a high speed polar code decoder using finite length scaling LDPC codes. J Adv Res Dyn Control Syst 10(10):153–161
12. Anuradha P, Arabelli R, Rajkumar K, Padmaja C (2022) MIMO-OFDM based high-rate information transmission through ultrasonic. In: AIP conference proceedings, vol 2418, no 1. AIP Publishing LLC, p 030059
13. Anuradha P, Rajkumar K, Arabelli R, Shareena R (2022) Implementation of high speed 64-bit MAC unit using FPGA. In: AIP conference proceedings, vol 2418, no 1. AIP Publishing LLC, p 030060
14. Rajkumar K, Anuradha P, Ravichander J, Vasavi J (2022) Cell-edge detection using canny algorithm and mathematical morphology. In: AIP conference proceedings, vol 2418, no 1. AIP Publishing LLC, p 030071

Automatic Vehicle Signal Enabling System

Sirisha Alamanda, Prathima Tirumalareddy, and Praneeth Prasad Vemuluri

Abstract Automatic signal enabling system is to make the indicators of the vehicle turn automatically on or off based on the predefined locations in maps in order to prevent the accidents which are prone only due to negligence in following few traffic rules, due to the laziness shown by the drivers in order to indicate the signal lights while shifting of lanes or turning on roads etc. In order to this, the proposed system use UV sensor to detect the turn beforehand and turn on LED's as alternative to signal indicators in vehicles.

Keywords Signal indicators · UV sensors · Raspberry pi 3 · Automatic vehicle signal enabling system · IoT · Mapbox GL ecosystem

1 Introduction

Every year many people die or many more are injured or disabled due to road accidents. The accidents occur due to the negligence of drivers in following the road rules or not using the indicators or headlights, etc. Many people have solved the issue of indicators of vehicle by making them automatic. The proposed automatic vehicle signal indicating system include maps for user to enter his source and destination which in turn gives the optimized directions for the user. The only thing the user needs to follow is the route given by the maps. The prototype of the proposed system uses LEDs which are alternatives for Indicators of vehicles and UV sensors for detecting the turn and uses MapBox directions API for getting the optimized directions. For every turn, the map gives us the direction along with the distance to the turn from the starting point (source) or from the turn. Taking the pre-defined average speed value for street as 30 kmph, the proposed system convert the distance into time to reach the turn which will in turn turns the indicators based on the time

S. Alamanda (✉) · P. Tirumalareddy · P. P. Vemuluri
Department of Information Technology, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: asirisha_it@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_8

provided. Once the turn is detected the indicator will be turned on and will continue to be on till the turn is taken then the indicator will be turned off after 5 seconds. The proposed system is also extended to work during lane switching.

2 Problem Definition

One major problem with many of the drivers is that they don't utilize the indicators of the vehicle which confuses the drivers behind the current vehicle. In many instances, particularly drivers of two-wheelers and less experienced drivers neglect the operation of indicators while taking turns. There are many vehicles signaling systems available in the market such as voice enabled or ORVM indicators etc., which work manually and need driver action to indicate the signals of the vehicles. Hence, the authors have proposed an automatic signal enabling system which makes the indicators of the vehicle turn on or off automatically without the intervention of the driver.

3 Related Work

Most of the existing turn indicator systems present in the vehicles are fully manually controlled systems. A semi-automatic signal indicator for two wheelers is proposed in [1]. Moving further there are few more automatic systems [2–6]. Voice detection is one of the methods in automatic systems [7]. Diving deep further into automatic systems [8], the authors have found that there are no such Automatic Vehicle Signal Enabling System which uses maps. The manual system is slow because the driver needs to make it work and in these systems the driver may forget to use the indicators. Coming to the voice-based indicators the voice-based module will just not hear the directions but every word of the driver, if there are more than one person and if they are having a conversation, a word of left or right may cause the signaling system to malfunction. And every moment the voice module hears the voice and processes the word it hears and makes it more complex and time taking to resolve.

4 Design of Proposed System

The proposed system mainly consists of two tiers. One is software tier and the other is hardware tier. The software tier is responsible of taking the user input, that is, the destination addresses and generating the directions provided by the route map.

4.1 Software Tier

The main aim of this tier is to get the directions from Java script library and rest API service and crosscheck between them and give us the optimized directions tuple.

4.2 Hardware Tier

The workflow of Hardware Tier is shown in Fig. 1, and the main aim of the hardware tier is to convert the directions tuple and use Raspberry Pi for turning the indicators on or off by using UV sensors.

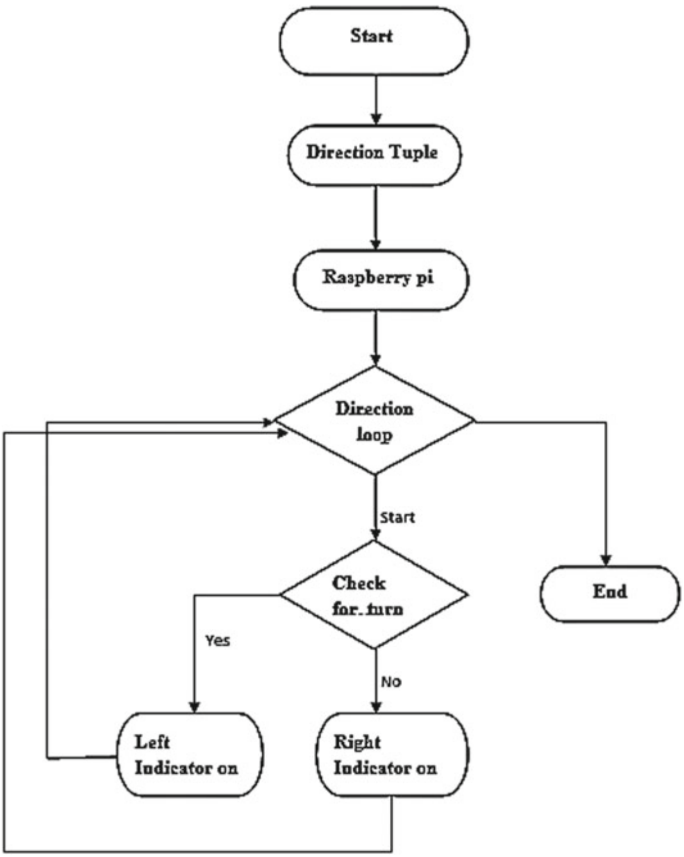


Fig. 1 Workflow of hardware tier

4.3 Java Script Libraries Used

Mapbox is a JavaScript library for interactive, customizable vector maps on the web. It takes map styles that conform to the Mapbox Style Specification, applies them to vector tiles that conform to the Mapbox Vector Tile Specification and renders them using WebGL. Mapbox GL JS is part of the cross-platform Mapbox GL ecosystem, which also includes compatible native SDKs for applications on Android, iOS, macOS, Qt, and React Native. Mapbox provides building blocks to add location features like maps, search, and navigation into any experience you create.

5 Methodology of the Proposed System

The proposed system has 4 major modules, they include software for user interface, integration of user interface with raspberry pi, hardware setup for turning on/off the LED's, cross checking with the directions provided by Java script library. The workflow of the proposed system is shown in Fig. 2.

5.1 Software for User Interface

This is where the user enters source/current and destination point. To implement this, the authors have used mapbox javascript library for displaying the maps and getting the directions for given source and destination. Now the user has to select from different options for his transit (car for instance) which gives us different directions. To handle these types of scenarios, there are different functions to handle them. One such case is driving by avoiding highways ignoring tollbooths gives one type of directions. Now the directions are shown to the users with the path in the map. These directions are transferred to backend or to raspberry pi once the user clicks on start navigation.

Fig. 2 Workflow of the proposed system

5.2 Integration of User Interface with Hardware

To integrate with hardware a code is written. This code is for detecting turns both left and right and turning on/off the LED's. Some such cases are as follows:

(a) What if the driver didn't take turn as suggested by the application?
(b) What if there is a road blockage in middle of the navigation path?
(c) What if there is huge traffic?

5.3 Hardware (Raspberry Pi3)

This part handles the working of UV sensors and LEDs by giving them the voltage required and plays an important role where it detects the turn by the rotation of the handle and turns the indicators on/off.

5.4 Cross Checking with Directions

As authors cannot depend on the java script library for directions for optimized route, so, it also cross checks them with other maps through their rest services. We provide the source and destination to the rest services with the options given by the user. The output of the rest services could be both json and xml format. Now it uses web scraping to parse the output of the json and xml format in the output format of java script and check which gives us the better optimized path from source to destination. For doing this formatting into directions tuple the authors have used beautiful soup.

6 Experimental Results and Discussions

The efficiency of the proposed system is checked by conducting some experiments on real time maps. Figure 3 shows the hardware setup of the proposed system which includes UV sensors and LED's.

First the user enters source/current and destination point using the user interface then the directions of the path are shown to the user as in Fig. 4.

These directions are converted into tokens/instructions and are retrieved to the Raspberry Pi as in Fig. 5. The Raspberry Pi then causes the vehicle signal indicators to turn on and off automatically as shown in Fig. 6 according to the instructions. The proposed system also shows the responses in term of left on or right on as in Fig. 7 when detected with a turn.

The authors' focus is to avoid the accidents that take place because of people forgetting the signal indicators. So, the proposed system has successfully addressed

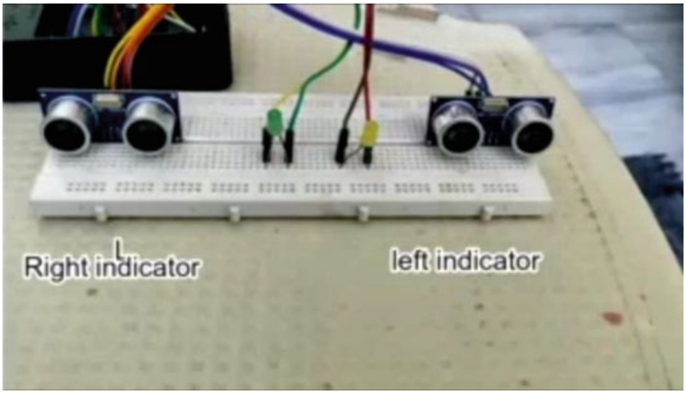


Fig. 3 Hardware showing UV sensors and LED's

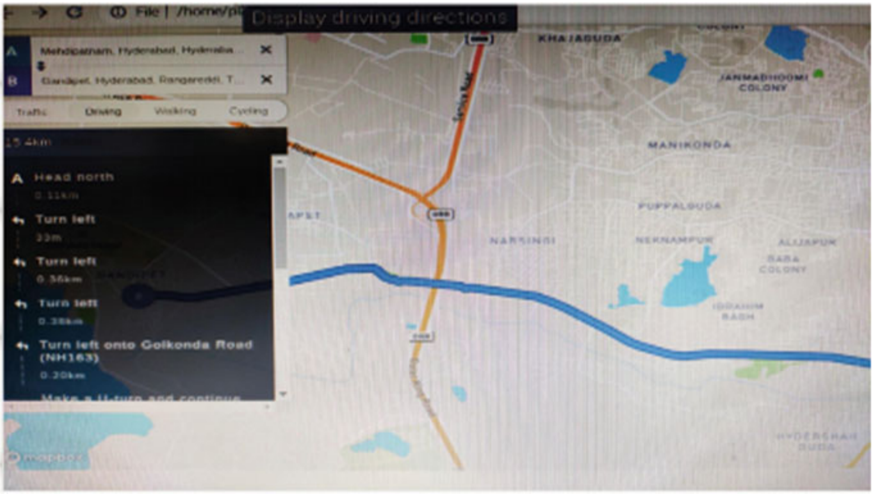


Fig. 4 Path directions from source to destination

the drawbacks of not using the indicators by making the indicators of the vehicle to turn on or turn off automatically as per the pre detected turns in the route of the given destination of the driver. The proposed system also works in the scenario where the user does not provide the destination. In such a scenario, when a user is taking the turn, the indicator of the vehicle turns on automatically if the user did not turn the indicator.

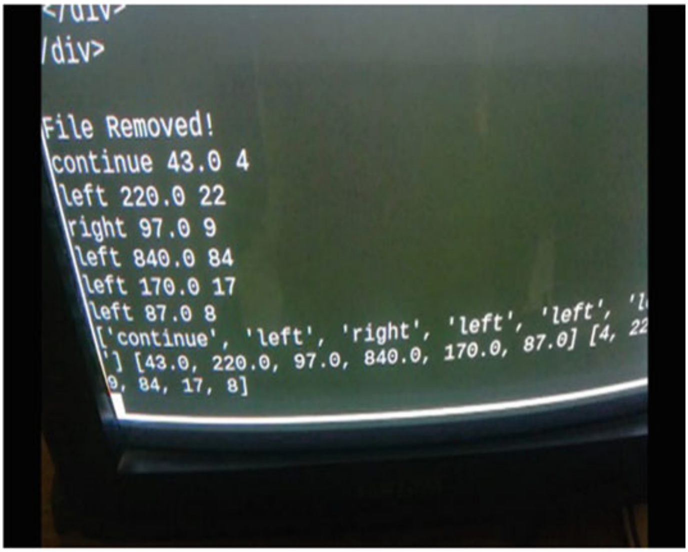


Fig. 5 Directions retrieved to Raspberry pi

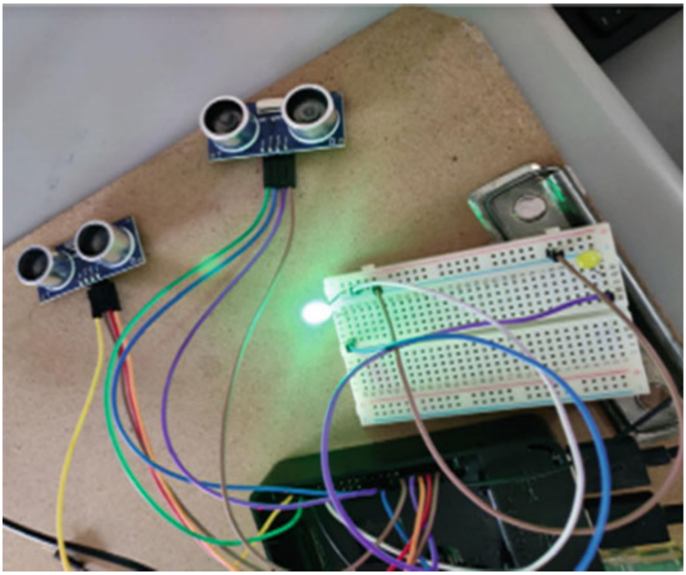


Fig. 6 Automatically signal indicator on when it detected a turn ahead

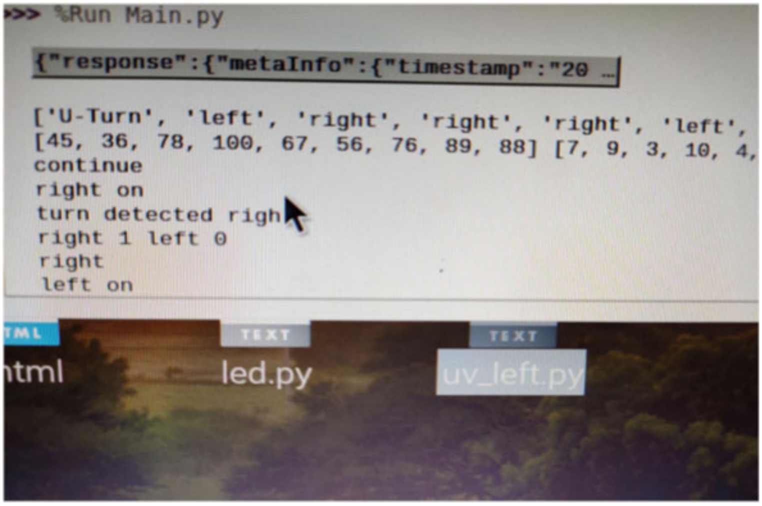


Fig. 7 When detected a turn, Raspberry pi returns signals

6.1 Applications

(a) Lane switching: There are some lanes on the national highways with speed limits defined for every lane. If we want to change from one lane to another lane, the indicator should be on. In this situation the proposed system can is used, no need to signal manually. Based on steering angle and based on route map, left/right side indicator will be ON automatically.

(b) Ambulances: Ambulances are one of the emergency vehicles, in which driver will take the patients and drop at the destination point. By using this Automatic Vehicle Signal Enabling System, the ambulance driver will enter destination from current location and he does not need to press the indicator button on/off manually. Automatically indicators will be on/off based on predefined locations in map. It is very useful for ambulances.

(c) Defense vehicles: The defense vehicles of the country should follow a route which must follow all the traffic rules. In this case, the vehicle must move at extreme right or extreme left.

(d) Autonomous vehicles: One of the best examples of the use of the proposed Automatic Vehicle Signal Enabling System is autonomous cars. And while fully autonomous cars are undoubtedly still a few years away, several automakers are rolling out semi-autonomous systems that are getting ever closer to the idea of jumping inside the car and letting it take over from there. In these vehicles also will develop automatic indicators ON/OFF.

7 Conclusion

The proposed system of automatic vehicle enabling overcomes the disadvantages of both manual type of indicators and voice-based indicators. The drawback of the proposed system is the need for internet connectivity and there is a compulsory need of a display device to take the destination input from the driver which adds more cost to the proposed system. The ultraviolet sensors need to be checked every two weeks. UV sensors are used for shorter distance detection only.

The proposed system can be extended to a display-less device where the driver enters the destination part in his mobile and can be used for lane switching. Though the authors have used raspberry pi3 in the proposed solution, a dedicated PCB can be used for the system which effectively lowers the price.

References

1. Divakara Murthy K, Harshitha V, Pedapudi A Design of semi-automatic signal indicator for a two-wheeler. St Peter's Engineering College, Mechanical Engineering Department, Hyderabad. ISSN (Print): 2393-8374
2. Rakesh S, Hegde NP (2021) Real-time area-based traffic density calculation using image processing for smart traffic system. In: Satapathy SC, Bhateja V, Favorskaya MN, Adilakshmi T (eds) Smart computing techniques and applications. Smart innovation, systems and technologies, vol 225. Springer, Singapore. https://doi.org/10.1007/978-981-16-0878-0_28
3. Sugamya K, Srinikitha N, Spandana C (2021) A dynamic density based traffic control system using IoT. In: 2nd international conference on IoT based control networks and intelligent systems (ICICNIS 2021), Proceedings of the international conference on IoT based control networks and intelligent systems—ICICNIS 2021. Available at SSRN: https://ssrn.com/abstract=3883885 or https://doi.org/10.2139/ssrn.3883885
4. Narayandas V, Dugyala R, Tiruvayipati S, Yellasiri R (2019) Necessity of MANET implementation over Internet of Things: the future of dynamic communication among end devices. In: Fifth international conference on image information processing (ICIIP). pp 359–362. https://doi.org/10.1109/ICIIP47207.2019.8985804
5. Vignesh M, Dhanashekaran V, Hemanth Kumar Reddy V, Neelesh, Preethi V (2018) Automatic vehicle turning indicator. SRM Institute of Science and Technology, Chennai
6. Perlman RI, Norman G (2004) Looper: wireless remote vehicle signal indicator for supplementing existing vehicle signal indicators. Patent Appl. No: US 6,677,856 B2
7. Raiyan Z, Nawaz MS, Adnan AKMA, Imam MH (2017) Design of an Arduino based voice-controlled automated wheelchair. In: IEEE region 10 humanitarian technology conference (R10-HTC). https://doi.org/10.1109/r10-htc.2017.8288954
8. Pagarkar MH, Parekh K, Shah J, Desai J, Advani P, Nikhil SS Automate vehicle control system

Study Report of Tor Antiforensic Techniques

Preeti S. Joshi and H. A. Dinesha

Abstract Internet has facilitated ease of communication, providing e-commerce, government and private services, also it has fostered crime to a large extent. Tor browser provides access to the dark web that is famous for market of illicit goods like drugs, weapons, pornography, malicious codes and hacked account details to name a few. All this is possible due to the anonymity and privacy provided by Tor. The anonymity and privacy provided by Tor network has made it a choice for cybercriminals. To avoid detection Tor uses antiforensic techniques like the browser settings, encryption, onion routing, onion addresses. The objective of this chapter is to put forth the antiforensic techniques provided by Tor network, present a detailed analysis of deanonymizing techniques researched till now and propose a simplified model for detecting a suspect using Tor communication.

Keywords Cybercrime · Tor · Anti-forensic · Deanonymization

1 Cybercrime

Cybercrime is a criminal or illegal activity that uses computer, network or networked devices. It is classified as those where computer is a target or a victim. Cybercrime can include major criminal activity like ransomware attacks, financial frauds like stealing or modifying financial or sensitive information etc. It may include activities that are unlawful as per the laws defined by the country. The usage of internet to perform day to day transactions has aided in increasing the volume and pace of cybercrime activities, as the criminal no longer needs to be physically present when committing a crime. The law enforcement agency faces difficulties since solving

P. S. Joshi (✉)
Department of Information Technology, Marathwada Mitramandal's College of Engineering, Pune, India
e-mail: preetijoshi@mmcoe.edu.in

H. A. Dinesha
Shridevi Institute of Engineering & Technology, Tumkur, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_9

such cases requires crossing national boundaries. The Internet's speed, convenience, anonymity are few other important facilitators for cybercrime. In this digital world, cybercrime committed by a person leaves clues of their identity, location of crime, activity, time, despite criminal's best efforts to cover the tracks. Digital forensic helps to find traces of events and evidences of activities on a digital device of a cybercrime. Anti-forensic techniques are major challenge to digital forensic investigators. Many anti-forensic techniques work by overwriting or erasing data, using cryptography to make data unintelligent or implement steganography and hide the data by embedding it in audio, video or image file. Some antiforensic techniques work by minimizing the attacker's footprint or attack the forensic tools [1].

There are four primary goals for antiforensics [2]:

1. Avoiding detection that some kind of event has taken place.
2. Disrupting the collection of information.
3. Increasing the time that an examiner needs to spend on a case.
4. Casting doubt on a forensic report or testimony

Of these primary goals mentioned in [2], Tor network attempts to satisfy first three objectives to a large extent. Communication using the Tor network is intended to increase the amount of time required by forensic examiners while also gaining privacy, anonymity, and avoiding detection. The anonymity and privacy provided by Tor network has made it a choice for cybercriminals. The services on Tor network range from simple communication services that facilitate anonymous communication (also used by journalists and whistle-blowers) with other parties to onion sites that sell illegal objects like drugs, weapons, credit card details etc. and sites that deal in child pornography, drug sale, human trafficking, terrorists and illegal sale of weapons, ransomware, bots, and PII to name a few. Numerous antiforensic methods are made available by Tor to promote illegal activity. It would be worthwhile to examine Tor usage on the network to commit a cybercrime or visit restricted websites (onion sites) and then find traces and proof of actions on the suspicious device, if any user is engaged in any unlawful or criminal behavior.

Major contributions of this paper are

1. To put forth the anti-forensic techniques provided by Tor network
2. Present a detailed analysis of deanonymizing techniques for Tor network
3. Propose a simplified model for detecting a suspect using Tor communication

2 Antiforensic Services on TOR

A Proxy server is a server on the Internet that helps to relay user's traffic. All traffic enters and leaves through the same server. The problem with simple proxy providers is that they create a single point of failure. The provider knows the client device, what is being accessed and can see traffic as it passes through their server. Tor is an overlay network onto the Internet, also called as darknet or dark-web. This part of the network can only be accessed through the Tor browser. Tor is the choice for criminal activities as it provides anonymity and privacy to the user. To reach the destination server, the communication across the Tor network is routed through three arbitrary relays and is encrypted at many stages. It's different from a typical proxy. In Tor network, the three relays from client to server are called entry or guard node, middle node and exit node. Each of the three relays has its own layer of encryption, each layer of encryption is decrypted by the relay onto the path so it's also called as "The onion routing". Each node only knows about its immediate predecessor and successor node and removes one layer of encryption. Data from exit node to the final destination are unencrypted and the server can see that its coming from IP address of the exit node. In this process of providing anonymity, Tor becomes slow network for end user. All sites available on Tor network fall under the "Onion" domain and are called as onion services or hidden services. For this, onion services IP address are not meaningful name resolution is not supported by DNS, instead there exist directory servers to handle the task of resolving addresses [3].

2.1 TOR Browser

The art of web accessible using regular browser like chrome, Firefox, Edge etc. is called surface web. A large portion of the web that is only accessible through Tor browser is called dark web. Like the incognito or private browsing mode of chrome and Firefox respectively, Tor browser does not store history of browsed sites, deletes browser cookies and login information. More over to use the Tor browser it is not necessary to install it on the device. Tails is a debian-based linux distribution that boots on live USB and connects to internet through Tor leaving no digital footprint. This satisfies the first goal of avoiding detection that some kind of event has taken place.

2.2 Encryption

Tor network connections are encrypted so any one monitoring the network cannot see what is being transmitted. The data is encrypted till it exits the third exit relay in the path. This satisfies the second goal of disrupting the collection of information and third goal of increasing the time that an examiner needs to spend on a case.

2.3 Onion Routing

When accessing any website using Tor, the server does not know the real IP of the client. The data is relayed through three relays working as proxies on Tor network and encrypted three times as it passes over these Tor relays. Due to the proxying mechanism, the local administrator does not detect real destination of data transmitted. This allows the user to visit sites blocked by your local administrator. This again satisfies the first goal of avoiding detection that some kind of event (connection to forbidden sites) has taken place.

3 Analysis of Deanonymization Techniques

3.1 Traffic Analysis

Security measures are taken to identify the traffic as Tor or non-Tor, these include flow analysis, correlation analysis, blocking based on IP address, port address and deep packet inspection. The identification based on flow analysis tries to identify patterns in flow of data and those in correlation attack try to correlate based on timing of the request/reply message. These implementation work on machine learning techniques to identify Tor connection. Various methods have been researched to identify Tor traffic and application used over Tor using traffic analysis. The characteristics like inter packet time, packet size, port number, etc. are used for traffic analysis. Mix network proved to be effective in providing anonymity and QoS, however accuracy of correlation-based traffic analysis on mix networks can be high as 100% [4]. In [5] authors have analyzed the flow and implemented correlation-based attack to determine tor traffic. The flow features are analyzed using machine learning algorithms to identify tor traffic in [6–10]. Vulnerability of tor users is demonstrated by analyzing packet size in [11]. Tor uses bandwidth weighted selection of relays, the traffic characteristics, link fingerprint are used to identify a node in Tor network [12]. Further advancement of traffic flow analysis using correlation attack by deep learning algorithm CNN is implemented in [13]. In this the proposed method Deepcorr does not need to learn target destinations of circuits, it learns and correlate two ends of Tor connection (Table 1).

3.2 Bridge Discovery

The analysis in these research papers have shown that anonymity of Tor is a myth and that IP addresses involved in connecting to Tor network can be identified. The public address of all tor relays can be downloaded from [14], so access to the Tor network can be blocked by blocking all publicly listed tor relay. To overcome IP blocking, the network settings in Tor browser allows user to select a bridge (a proxy, not listed in the main Tor directory) to connect to Tor network. The addition of bridges to TOR is a step forward in providing anonymity in Tor networks. Bridges always act as first hop in a circuit, and whose IP addresses are not publicly advertised. Discovery Tor bridges by two approaches namely emails and https enumeration and by setting

Table 1 Tor traffic identification and analysis

Paper	Investigation	Method	Outcome	Tools
Shahbar and Zincir-Heywood [9]	Circuit level classification using cell attributes of Tor Flow level classification of Tor	Machine learning algorithms Naïve Bayes, Bayes Net, C4.5, random forest	100% accuracy for both circuit and flow level classification	Tranalyser2, Tcptrace, Weka
Barker et al. [11]	Classification of Tor traffic	Machine learning algorithms random forest, j4.8, Adaboost, packet size	93.7% accuracy for HTTP over Tor and 97.7% for HTTPS over Tor	Weka
Cuzzocrea et al. [6]	(a) Flow analysis to identify Tor traffic (b) Type of activity/service used over Tor	Machine learning algorithm J48, J48 consolidated, BayesNet, jRip, OneR, REPTree	Precision and recall = 1 for Tor traffic identification Precision and recall = 0.998 for service identification	Wireshark, tcpdump, ISCXFlowMeter
Habibi Lashkari et al. [7]	Time based feature to identify Tor traffic and application over Tor	Machine learning algorithms: C4.5, random forest, KNN, ZeroR	1. Identify Tor traffic 2. Detect 8 applications over Tor: browsing, email, chat, audio streaming, video streaming, File transfer, P2P, VoIP	ISCXFlowMeter
Almubayed et al. [16]	Identifying Tor usage in an offline network trace via websites fingerprinting	Machine learning: Naïve Bayes, decision tree: C4.5, random forest, SVM	Accuracy of 99.64%	NetMate for flow generation
Montieri et al. [10]	Traffic analysis using machine learning to identify Tor, I2P and Jon Donym and application	Machine learning algorithms: Naïve Bayes, decision tree (C4.5, random forest)	1. Accuracy of Tor traffic identification 99.87% 2. Application over Tor 73.99%	Tranalyser2, Dataset: Anon17
Saputra et al. [17]	Detect and block Tor traffic	Deep packet inspection	TLS handshake process and certificates identify Tor traffic	Wireshark, Bro-IDS
Chakravarty et al. [5]	Statistical correlation of traffic at client and controlled server	Pearson's correlation between server injected square wave and step pattern	Accuracy to detect Tor source 81.4%	NetFlow

(continued)

Table 1 (continued)

Paper	Investigation	Method	Outcome	Tools
Nasr et al. [13]	Flow correlation in Tor connections	CNN algorithm	Flow correlation accuracy of 96%	–

Table 2 TOR bridge discovery

Paper	Investigation	Method	Outcome
Ling et al. [15]	Identify IP address of bridges in Tor network	Sending bulk emails to request bridge address	2365 Tor bridges discovered
		Setting malicious middle router	2369 Tor bridges discovered
Matic et al. [18]	Security analysis of public and private bridge infrastructure	1. CollecTor service 2. Scan search engines Shodhan & Censys	694 private bridges and 1292 public bridges identified

malicious tor middle router is analyzed in [15] and shown that the latter is simple, efficient and effective with low overhead (Table 2).

3.3 Other Approaches

Other approaches to that prove to be a threat to anonymity of Tor are user tracking by click path analytics [19], and DNS leaks due to numerous. Onion requests observed at global DNS infrastructure [20]. A simplified procedure compared to the popular method of DPI is TLS certificate parameters, port number, the contents of the X.509 certificate name and its size are used to identify Tor connection [21] (Table 3).

3.4 Pluggable Transports

Access to the Tor network may sometimes be blocked by installing censors. An increasing number of censoring countries are using Deep Packet Inspection (DPI) to classify Internet traffic flows by protocol. Tor uses bridge relays to get around censors that block by IP address. Deep packet inspection to identify that the traffic carries Tor data is researched in [17]. A security analysis of public bridges using CollecTor service and private bridges with their infrastructure is discussed in [18].

In recent years, methods have been are developed to prevent Tor even when clients are using bridges. Typically, they accomplish this by scanning network traffic to find

Table 3 Other approaches to deanonymize TOR

Paper	Investigation	Method	Outcome
Mohaisen and Ren [20]	Measurement of leakage of. Onion domain names in DNS infrastructure	A and J root servers from DITL dataset	Increase in onion request at DNS server due to human error and DNS prefetching
Vlajic et al. [19]	User tracking from destination system	1. Tor browser and IP address-based tracking 2.Tor browser and cookie based 3.Tor browser and URL rewriting based 4.Tor browser and ETag based	Tor browser provides limited protection against 4 methods of tracking
Lapshichyov and Makarevich [21]	Identification and blocking of the connection to the Tor network	Analysis using the forensic network analysis tool NetworkMiner	Tor TLS certificates identified

Tor. The Tor browser comes with what are known as pluggable transport tools for getting around these obstacles. These transports alter all Tor traffic so that it cannot be recognized as a Tor connection between the client and its first hop. Censors monitoring the traffic between the client and the bridge instead see traffic that has been altered but still seems normal non-Tor traffic because pluggable transports are used to transform the Tor communication flow between the client and the bridge. Pluggable transports also called as obfuscators can be categorized in several types. Some of the well-known types provide randomizing, protocol mimicry, and tunneling [22]. The meek pluggable transport implements the concept of domain fronting, that is, it hides the target domain of a connection by obfuscating the SNI field of a TLS connection, it requires finding a hosting provider or content delivery network (CDN) which has a certificate that supports multiple target domains. The traffic looks like normal HTTPS traffic that to cloud platforms like Amazon, Akamai or Google and preventing access to such cloud platforms is not easy. Currently domain fronting is available only on Azure platform.

The Tor browser setting allows user to select a PT. Some of the PT's integrated in the browser are Obfs4, snowflakes, meek, FTE etc. But pluggable transports are not immune to detection, if a censor is given enough time. Robustness of PT's namely flashproxy, scramblesuit, FTE, Meek, obfs3 is investigated in [23] to conclude that these PT's have their own unique fingerprints that makes them identifiable. Meek does not mask traffic volume and is vulnerable to statistical attacks based on traffic volume and timing of sent and received messages or hidden Markov model as in [24] (Table 4).

Table 4 TOR traffic under pluggable transport

Paper	Pluggable transport	Detection method	Outcome
He et al. [26]	Obfs4	Two level filtering 1.Based on randomness and timing characteristics of Obfs4 2. SVM classification algorithms	Accuracy of SVM classifier is 99%
Shahbar and Zincir-Heywood [23]	FTE, obfs3, meek, ScarmbleSuit, flashproxy	Tranalyzer as flow exporter and C4.5 decision tree to pluggable transport	97% correct classification
Shahbar and Zincir-Heywood [27]	FTE, obfs3, meek, ScarmleSuit, flashproxy	Tranalyzer as flow exporter and C4.5 decision tree to pluggable transport	Classify Tor pluggable transports describing the proper features
Wang et al. [28]	Obfs3, obfs4, FTE, meek-amazon, meek-google	Fast entropy-based tests for randomizer protocols and FTE, and slightly less efficient machine learning-based attacks that reliably detect meek	Detect 100% of obfsproxy3/ obfsproxy4 and FTE traces and 98% for meek variants
Yao et al. [24]	Meek	Mixure of Gaussian based HMM, to characterize Inter-packet time and packet size	Identification rate of 99.4%
Berenjestanaki and Akhaee [29]	Tor, UltraSurf, and ScrambleSuit	Machine learning and flow analysis	Differentiate Tor, Ultrasurf and ScrambleSuit
Mohajeri Moghaddam et al. [30]	SkypMorph	Naive traffic shaping, source distribution	Morph Tor streams into traffic with indistinguishable packet sizes and timings to Skype video
Weinberg et al. [31]	StegoTorus	Provides a collection of steganography modules and implements a unique encryption technique	Makes Tor more resistant to fingerprint attacks and provides usable performance

4 Proposed Model

With all the antiforensic techniques provided by Tor Browser and Tor network, Tor is a convenient tool for cyber criminals. The anonymity and privacy foster cyber-crime. In a large enterprise network, it becomes challenging for an investigator to trace the device and the user accessing Tor network services. Finding devices that produce unlawful or malicious traffic is a key objective of network forensic techniques. Collecting, identifying, examining, correlating, analyzing, and documenting digital is required for this in order to discover facts. Organizations need to be aware of how well their network access rules and regulations have been protected and put in place. The environment should help any forensic investigator conducting an inquiry in the event of any cybercrime. We propose an easy to deploy identification method for cybercrime using Tor in a large-scale network. It is to analyze the traffic and find traces of Tor connection and identify the particular device that made a connection to Tor network. Network access made by client devices is routed through a monitoring device that captures outgoing traffic in PCAP files. The captured PCAP file is then loaded in Netresec's Caploader tool [25]. CapLoader implements a statistical method for protocol detection to identify Tor traffic. The source address connecting to Tor network from the egress traffic leads to the offender's device. Further the offender's device is investigated for traces of the activity by forensic analysis (Fig. 1).

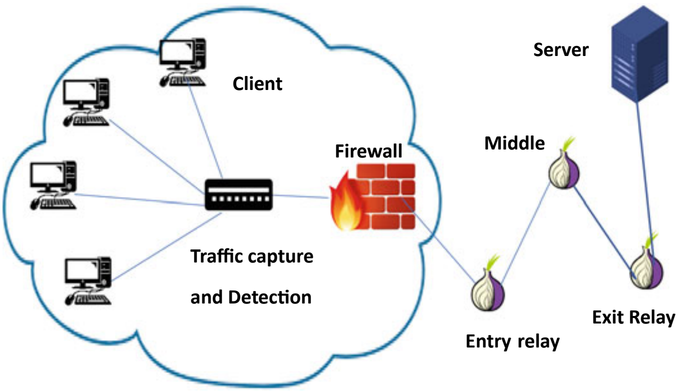


Fig. 1 Tor traffic detection to identify suspect in the large enterprise network

5 Conclusion

In this paper, we described the anti-forensic techniques of Tor that are supportive to criminal activities. We analyzed various techniques of deanonymizing Tor traffic explored by the researchers. Individual traffic flow analysis provides cybercrime investigators with valuable information. We proposed a model for detecting Tor connection from traffic in a large enterprise network using CapLoader tool and statistical method. This will help local network administrator to segregate Tor traffic and identify offender's device. The solution aids to satisfy and meet the law and regulations and recognize threats to reputations of the enterprise.

References

1. Garfinkel S (2007) Anti-forensics: techniques, detection and countermeasures. In: 2nd international conference in i-warefare and security. pp 77
2. Bleeding-edge anti-forensics. In: Infosec world conference and Expo. MIS Training Institute
3. https://support.torproject.org/about/
4. Zhu Y, Fu X, Graham B, Bettati R, Zhao W (2010) Correlation-based traffic analysis attacks on anonymity networks. IEEE Trans Parallel Distrib Syst 21(7):954–967
5. Chakravarty S, Barbera MV, Portokalidis G, Polychronakis M, Keromytis AD (2014) On the effectiveness of traffic analysis against anonymity networks using flow records. In: Faloutsos M, Kuzmanovic A (eds) Passive and active measurement. PAM. Lecture notes in computer science, vol 8362. Springer, Cham
6. Cuzzocrea A, Martinelli F, Mercaldo F, Vercelli G (2017) Tor traffic analysis and detection via machine learning techniques. In: IEEE international conference on big data (big data). pp 4474–4480
7. Habibi Lashkari A, Draper Gil G, Mamun M, Ghorbani A (2017) Characterization of tor traffic using time based features. In: Proceedings of the 3rd international conference on information systems security and privacy—ICISSP. pp 253–262. ISBN 978-989-758-209-7; ISSN 2184-4356
8. Mercaldo F, Martinelli F (2017) Tor traffic analysis and identification. In: AEIT international annual conference. pp 1–6
9. Shahbar K, Zincir-Heywood AN (2014) Benchmarking two techniques for tor classification: flow level and circuit level classification. In: IEEE symposium on computational intelligence in cyber security (CICS). pp 1–8
10. Montieri A, Ciuonzo D, Aceto G, Pescapé A (2018) Anonymity services tor, I2P, JonDonym: classifying in the dark (web). IEEE Trans Dependable Secure Comput. https://doi.org/10.1109/TDSC.2018.2804394
11. Barker J, Hannay P, Szewczyk P (2011) Using traffic analysis to identify the second generation onion router. In: IFIP 9th international conference on embedded and ubiquitous computing. pp 72–78
12. Shen S, Gao J, Wu A (2018) Weakness identification and flow analysis based on tor network. In: Proceedings of the 8th international conference on communication and network security (ICCNS 2018). Association for Computing Machinery, New York, NY, USA, pp 90–94
13. Nasr M, Bahramali A, Houmansadr A (2018) DeepCorr: strong flow correlation attacks on tor using deep learning. In: Proceedings of the 2018 ACM SIGSAC conference on computer and communications security (CCS '18). Association for Computing Machinery, New York, NY, USA, pp 1962–1976
14. https://www.dan.me.uk/tornodes. Last Accessed 21 Nov 2022
15. Ling Z, Luo J, Yu W, Yang M, Fu X (2015) Tor bridge discovery: extensive analysis and large-scale empirical evaluation. IEEE Trans Parallel Distrib Syst 26(07):1887–1899
16. Almubayed A, Hadi A, Atoum J (2015) A model for detecting tor encrypted traffic using supervised machine learning. Int J Comput Network Inf Secur. https://doi.org/10.5815/ijcnis.2015.07.02
17. Saputra FA, Nadhori IU, Barry BF (2016) Detecting and blocking onion router traffic using deep packet inspection. In: International electronics symposium (IES). pp 283–288
18. Matic S, Troncoso C, Caballero J (2017) Dissecting tor bridges: a security evaluation of their private and public infrastructures. https://doi.org/10.14722/ndss.2017.23345
19. Vlajic N, Madani P, Nguyen E (2017) Anonymity of tor users demystified. In: International conference on computational science and computational intelligence (CSCI). pp 109–114
20. Mohaisen A, Ren K (2017) Leakage of. onion at the DNS root: measurements, causes, and countermeasures. IEEE/ACM Trans Networking 25(5):3059–3072
21. Lapshichyov V, Makarevich O (2019) TLS certificate as a sign of establishing a connection with the network tor. In: Proceedings of the 12th international conference on security of information

and networks (SIN '19). Association for Computing Machinery, New York, NY, USA, Article 14, pp 1–6
22. Dixon L, Ristenpart T, Shrimpton T (2016) Network traffic obfuscation and automated internet censorship. In: IEEE security and privacy, vol 14, no 6. pp 43–53
23. Shahbar K, Zincir-Heywood AN (2015) Traffic flow analysis of tor pluggable transports. In: 11th international conference on network and service management (CNSM), Barcelona, pp 178–181
24. Yao Z et al (2018) Meek-based tor traffic identification with hidden Markov model. In: IEEE 20th international conference on high performance computing and communications; IEEE 16th international conference on smart city; IEEE 4th international conference on data science and systems (HPCC/SmartCity/DSS), Exeter, United Kingdom, pp 335–340
25. https://netresec.com/?b=13400BE. Last accessed 21 Nov 2022
26. He Y, Hu L, Gao R (2019) Detection of tor traffic hiding under Obfs4 protocol based on two-level filtering. In: 2nd international conference on data intelligence and security (ICDIS), South Padre Island, TX, USA, pp 195–200
27. Shahbar K, Zincir-Heywood AN An analysis of tor pluggable transports under adversarial conditions. In: IEEE symposium series on computational intelligence (SSCI), Honolulu, HI. pp 1–7
28. Wang L, Dyer KP, Akella A, Ristenpart T, Shrimpton T (2015) Seeing through network-protocol obfuscation. In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security (CCS '15). Association for Computing Machinery, New York, NY, USA, pp 57–69
29. Berenjestanaki MH, Akhaee M (2018) Application detection in anonymous communication networks. In: Proceedings of the central European cybersecurity conference. Association for Computing Machinery, New York, NY, USA, Article 12, pp 1–5
30. Mohajeri Moghaddam H, Li B, Derakhshani M, Goldberg I (2012) SkypeMorph: protocol obfuscation for tor bridges. In: Proceedings of the 2012 ACM conference on computer and communications security (CCS '12). Association for Computing Machinery, New York, NY, USA, pp 97–108
31. Weinberg Z, Wang J, Yegneswaran V, Briesemeister L, Cheung S, Wang F, Boneh D (2012) StegoTorus: a camouflage proxy for the tor anonymity system. In: Proceedings of the ACM conference on computer and communications security (CCS '12). Association for Computing Machinery, New York, NY, USA, pp 109–120

Analysis of Underlying and Forecasting Factors of Type 1 Diabetes and Prediction of Diabetes Using Machine Learning

K. Poorani and M. Karuppasamy

Abstract Diabetes mellitus is a chronic metabolic syndrome. This is due to high glucose levels in the blood. The glucose has to be broken down as energy with the hormone called insulin. There are situations where insulin is deficient or resistant which leads to spike in blood glucose levels. Type 1 diabetes (T1D), also known as insulin deficient and type 2 (T2D) are called insulin resistant. Since type 1 generally prevails due to various factors and affects people of all age, this proposed work concentrates on the underlying modifiable and nonmodifiable factors responsible for the causes of type 1 diabetes. This work also extends the forecast analysis on complications that may arise due to longstanding type 1 diabetes. In order to find diabetes at the earlier stage the machine learning algorithm has been implemented. kNN algorithm has been implemented with certain modification to present an efficient model for classification.

Keywords Diabetes mellitus · Factors influencing diabetes · Forecast analysis · Prediction · kNN model

1 Introduction

Type 1 diabetes is an auto immune disorder prevailing worldwide [1]. T1D is also caused by autoimmune reaction, which destroys the cells in the pancreas. This causes destruction in the beta cells of the pancreas, which makes insulin. Insulin is a hormone which controls the blood glucose levels.

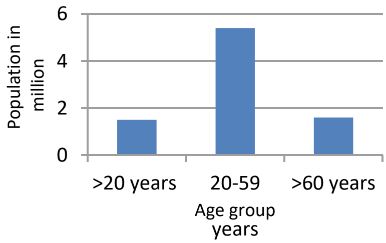
Gregory et al. [2] made an analysis on the prevalence and mortality rate regarding Type 1 diabetes mellitus. Their findings convey that in 2021, there were about

K. Poorani (✉) · M. Karuppasamy
Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India
e-mail: pooranikhari@gmail.com

M. Karuppasamy
e-mail: karuppasamy.m1987@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_10

Fig. 1 Prevalence of type 1 diabetes worldwide



8.4 million individuals worldwide with type 1 diabetes. Their findings has been described in Fig. 1. The cases are more around the age of 20–59.

The most common reasons for diabetes are listed here. Some of the important causes are less knowledge and education regarding the disease, the high intake of carbohydrate food, low consumption of fruits, nuts, seeds, whole grains, less exercise (Physically inactive), tobacco and alcohol intake, environmental pollution, high blood pressure and cholesterol are the various underlying causes of diabetes of any type. Zóka et al. [3] examines that genetics has a role in transferring genes from one generation to the other. They also examined the beta-cell functionalities and immune responses that causes and viral infections and certain environmental conditions also pave the way for onset of T1D.

The root cause and the factors causing type 1 and type 2 varies [4]. Obesity is the main cause of type 1 or type 2 diabetes which is observed globally at various demographic region. Since children who are obese tend to be less active physically but consume more which automatically paves way for the onset of diabetes. Corbin et al. [5] in his paper presents the impact of obesity and diabetes. People who are obese tend to be more inactive than with normal people.

Type 1 diabetes is a disease of childhood, which affects the children rather than type 2, which affects adults the most. In type 1 diabetes, insulin productivity is not found, hence leading to insulin dependant. Type 1 is not only found during childhood but in adult stages also. Blood glucose level should be maintained at appropriate levels to avoid complications.

Figure 2 represents the nonmodifiable and modifiable factors for the onset of Type 1 diabetes. The first in the list is the autoimmune disorder which may cause diabetes. Genetics also cause Type 1 diabetes. Obesity is also one of the major factors for the onset of diabetes of all types. There are also many drugs that are responsible for the onset of diabetes. The other factors include environmental factors, pancreas damage, tobacco use, smoking, alcohol intake and so on.

The causes can also be categorized as modifiable and nonmodifiable factors. Nonmodifiable factors cannot be changed. But modifiable factors like glycemic control, being physically inactive, blood pressure and cholesterol levels, obesity, smoking, alcohol usage, tobacco use can be prohibited. HbA1 C levels which are usually taken for diabetes with certain intervals (once in 3 months) needs to be in the exact range. So that blood glucose level can be maintained. However, taking medicine for controlling blood glucose maintains glucose levels, other modifiable factors need

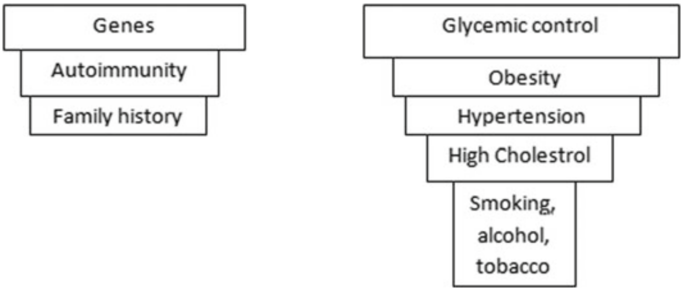


Fig. 2 Non-modifiable and modifiable factors of type 1 diabetes

to be maintained with great care so that complications can be avoided. Fluctuations in HbA1C levels leads to the onset of complications.

Type 1 diabetes leads to many complications in diabetic patients. If the blood glucose is maintained within the minimum range, then there is low probability for complications. If the blood glucose level is high and is not properly monitored, then there is a lot of chance for higher complications like heart disease, eye disease (retinopathy), nerve damage (neuropathy), and kidney damage (nephropathy), Alzheimer's disease, dementia and foot problems. As the cases of type 1 is growing the demand for insulin production is also increasing.

Perkins et al. [6] highlights that the kidney damage is associated with the type 1 diabetes. Whenever diabetes is not maintained properly there causes high blood glucose level which automatically destroys the small and large blood vessels leading to kidney damage. Not only kidney, it tends to affect various parts of the internal organs with much less symptoms. Thus blood glucose levels should be monitored regularly.

Gubitosi-Klug et al. [7] examined the association of micro vascular complication with the risk of cardiovascular diseases (CVD). Their findings prove that the micro vascular complications like retinopathy and nephropathy leads to risk of CVD. Their results show that the glomerular filter rate < 60 mL/min or cases with severe albuminuria are highly prone to CVD. Since type 1 are insulin dependent, people with type 1 diabetes are almost dependent on the insulin usage. In order to have insulin at proper levels, the monitoring of blood glucose levels becomes a necessity.

Steigmann et al. [8] has done an real-time study with type 1 people participated in epidemiology of diabetes intervention and complication study and found that the prevalence of oral disease is found with type 1 diabetes. It was found that the presence of long-standing diabetes with type 1 is associated with presence of periodontitis and poor cholesterol level. The poor glycemic control also added to the complication list. Chiesa and Marcovecchio [9] formulated the risk assessment factors with cardio metabolic factors and their role in onset of cardiovascular diseases. Even though HbA1c is the main target for diabetic people glycemic control is not well maintained in people with Type 1. Some early intervention methods like urine albumin analysis can help to reduce the complication of cardiovascular diseases since childhood diabetes have more probability of developing cardiovascular diseases.

2 Related Work

Shuja et al. [10] explained the prediction of diabetes with the help of data mining classifiers and also implemented SMOTE method to provide better results. However, it gives good results nowadays machine learning algorithms has been implemented to find the classification. Hasan et al. [11] in their findings employed various machine learning algorithms like kNN, decision tree, random forest, Ada boost, Naïve bayes, XGBoost and multilayer perceptron and gave better results when used with weighted ensembling method. The implementation of ensembling technique has been utilized to give good results in terms of accuracy, specificity and sensitivity.

Dutta et al. [12] explained the importance of feature selection in the dataset so that the accuracy of the model can be well established. Machine learning algorithms and found that knn with wrapper algorithm for feature selection gave high accuracy. Since the training set and test set has not been defined with any metrics.

Sneha and Ganglin [13] examined the analysis of diabetes features using optimal feature selection method using machine learning algorithm. They proposed decision tree, naïve bayes, kNN, random forest and other machine learning algorithms and found naïve bayes to be better model in terms of accuracy.

Analysis of all the above work brings this proposed system which implements the KNN classifier for classification. kNN works well with both classification and regression. This model also reduces the need to tune the parameters for classification but modified approach can be done. kNN model is easy to perform interpretation and also it is less time consuming, which is the major factor considered for building a good model.

3 Proposed Method

The machine learning algorithm has been implemented for the classification of diabetes. Pima Indian Diabetes Dataset (PIDD) has been utilized for the processing. The data are preprocessed before implementing the machine learning algorithm so that it gives better results in terms of accuracy and sensitivity.

Pima Indian Diabetic Dataset (PIDD) has been used to perform classification of diabetes. Since this dataset utilizes all the factors to be analyzed for the prediction of diabetes, every features are taken for the prediction. Since the outcome is a labeled data it is easy for performing classification using supervised machine learning techniques. Dataset used for the prediction of diabetes is described in Table 1.

Table 1 PIMA IDD Dataset description

S. No.	Attribute description	Description of attributes
1	Pregnancy	Number of pregnancy
2	Glucose level	Plasma glucose level 2-h OGTT
3	Blood pressure range (BP)	BP range
4	Skin thickness measure	Skin foldness using collagen content
5	Serum insulin	2 serum insulin level
6	BMI	Body mass index (Kg/m^2)
7	Diabetes pedigree function	Attribute that shows family history
8	Age	Age
9	Outcome	Diabetes class variable, 0-diabetic 1-not diabetic

Preprocessing the dataset has been done with normalization method called min–max normalization.

$$\mathrm{x}' = \mathrm{x} - \min(\mathrm{x})/\max(\mathrm{x}) - \min(\mathrm{x})$$

Figure 3 depicts the proposed model for the prediction of diabetes in terms of machine learning flow. The stage includes preprocessing using min max scaling, then data have been splitted for training and testing with random sampling method. K-fold validation is commonly used method to select a classifier and also to examine the performance.

Figure 4 represents the tenfold cross validation process for the Pima Indians diabetic dataset, which is the most commonly used dataset for classification. Various machine learning algorithm has been used to classify the diabetes. Several machine learning algorithms have been employed to find the onset of diabetes in the earlier stage. Among all the algorithms, kNN seems to find the best predictor in terms of sensitivity. kNN seems to be the best choice for the classification of diabetes dataset since the size of the dataset is not very large. Before the classification process it is a must to preprocess the data. kNN algorithm is being implemented to classify the

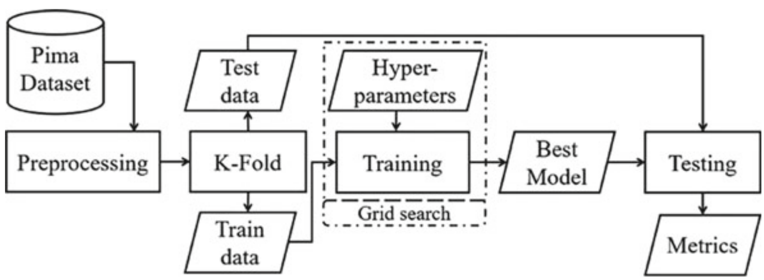


Fig. 3 Proposed model for the prediction of diabetes

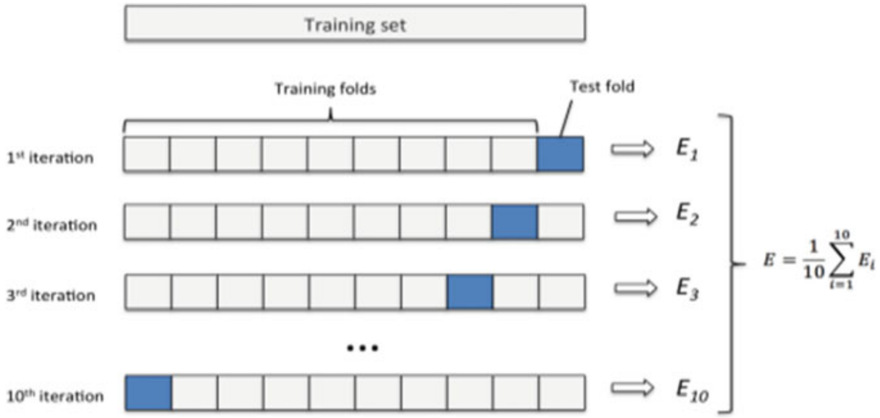


Fig. 4 10-fold cross validation

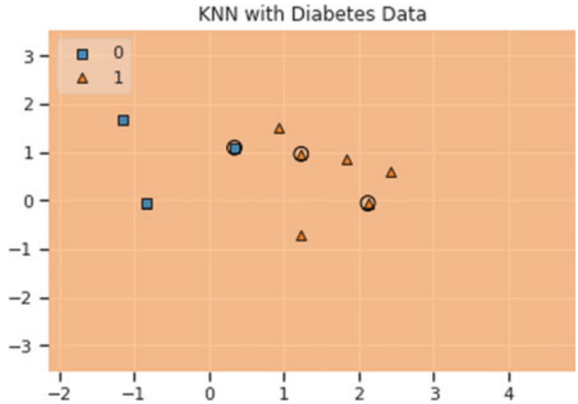
diabetes and nondiabetes cases. Scaling has been performed to bring all the data to the same scale so that distance measurement using Euclidean distance will be easier.

KNearest Neighbor is a supervised machine learning algorithm. Supervised machine learning is used in cases of classification and regression. It is used to find new pattern from the already existing patterns. It is used to classify an object using k training data. The k can take values more than one.

Figure 5 depicts the classification of diabetes data using kNN model. This is easier way of representing data. The 0 represents non-diabetic cases and 1 represents diabetic cases. The below Fig. 6 represents confusion matrix obtained after implementing the kNN algorithm.

Confusion matrix is one of the metrics used to evaluate the accuracy of the model. Thus by the above confusion matrix the accuracy of the model is 77%. Since the model seems to be less accurate it has been enhanced with the random sampling method so that every kind of data will be given chance to involve in training. The

Fig. 5 Sample model of data classification using kNN



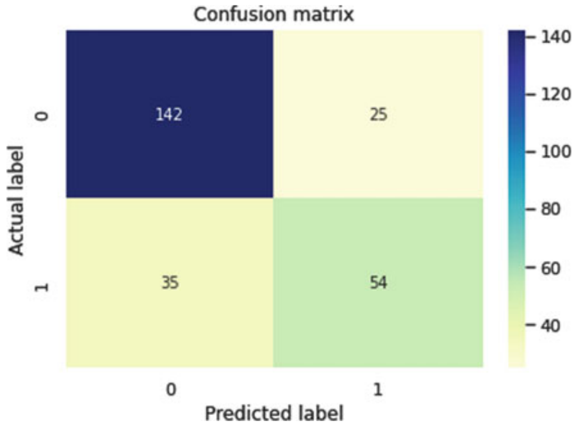


Fig. 6 Confusion matrix implementing kNN with pima

sensitivity (recall) of the proposed model is found to be 85% which gives a better satisfactory prediction model. The result gave an accuracy of 77%. The sensitivity of the model is found to be improved.

4 Conclusion and Future Discussion

The overall analysis shows that the type 1 diabetes is caused by various factors and leads to several complications. So monitoring and maintaining blood glucose level can help diabetic patient to reduce the complication risk associated with the disease. This work highlights the underlying causes of type 1 diabetes and its overall impact and suggests to maintain HbA1c levels accordingly. This works also extend the prediction of diabetes using modified kNN method to give better predictive model. In future ANN algorithm can be implemented to give a better prediction model.

References

1. Center for Disease Control and Prevention (CDC). https://www.cdc.gov/diabetes/basics/diabetes.html
2. Gregory GA, Robinson TI, Linklater SE, Wang F, Colagiuri S, de Beaufort C, Donaghue KC, Magliano DJ, Maniam J, Orchard TJ, Rai P (2022) Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. Lancet Diab Endocrinol 10(10):741–760
3. Zóka A, Műzes G, Somogyi A, Varga T, Szémán B, Al-Aissa Z, Hadarits O, Firneisz G (2013) Altered immune regulation in type 1 diabetes. Clin Dev Immunol
4. Mayo clinic. https://www.mayoclinic.org/diseases-conditions/diabetes/symptoms-causes/syc-20371444

5. Corbin KD, Driscoll KA, Pratley RE, Smith SR, Maahs DM, Mayer-Davis EJ (2018) Advancing care for type 1 diabetes and obesity network (ACT1ON). Obesity in type 1 diabetes: pathophysiology, clinical impact, and mechanisms. Endocr Rev 39(5):629–663
6. Perkins BA, Bebu I, de Boer IH, Molitch M, Tamborlane W, Lorenzi G, Herman W, White NH, Pop-Busui R, Paterson AD, Orchard T (2019) Risk factors for kidney disease in type 1 diabetes. Diabetes Care 42(5):883–890
7. Gubitosi-Klug R, Gao X, Pop-Busui R, de Boer IH, White N, Aiello LP, Miller R, Palmer J, Tamborlane W, Wallia A, Kosiborod M (2021) Associations of microvascular complications with the risk of cardiovascular disease in type 1 diabetes. Diabetes Care 44(7):1499–1505
8. Steigmann L, Miller R, Trapani VR, Giannobile WV, Braffett BH, Pop-Busui R, Lorenzi G, Herman WH, Sarma AV (2022) Diabetes control and complications trial, epidemiology of diabetes interventions and complications research group. Type 1 diabetes and oral health: findings from the epidemiology of diabetes interventions and complications (EDIC) study. J Diabetes Complications 36(4):108120
9. Chiesa ST, Marcovecchio ML (2021) Preventing cardiovascular complications in type 1 diabetes: the need for a lifetime approach. Front Pediatr 9:696499
10. Shuja M, Mittal S, Zaman M (2020) Effective prediction of type ii diabetes mellitus using data mining classifiers and SMOTE. In: Advances in computing and intelligent systems. Springer, Singapore, pp 195–211
11. Hasan MK, Alam MA, Das D, Hossain E, Hasan M (2020) Diabetes prediction using ensembling of different machine learning classifiers. IEEE Access 8:76516–76531
12. Dutta D, Paul D, Ghosh P (2018) Analysing feature importances for diabetes prediction using machine learning. In: IEEE 9th annual information technology, electronics and mobile communication conference (IEMCON). IEEE, pp 924–928
13. Sneha N, Gangil T (2019) Analysis of diabetes mellitus for early prediction using optimal features selection. J Big Data 6(1):1–9

An Enhancement in Crypto Key Generation Using Image Features with CRT

Kalyanapu Srinivas, Noothi Sravan Kumar, T. Sanathkumar, and K. Rama Devi

Abstract Strong cryptographic algorithms are the main sources for building security systems. However, providing security depends on the secret information like passwords, cryptographic keys, etc. In information technology, cryptography is the term representing data and secret information used to protect data that are vulnerable to unauthorized access during transmission. An algorithm and a cryptographic key are the basic components of cryptography. The key for cryptographic algorithms generated using traditional techniques are suffering from guessing, dictionary attacks, social engineering and shoulder surfing. This chapter points out the importance of key for any cryptographic algorithm. To enhance security, a new key generation technique is proposed using images with Chinese Remainder Theorem. The proposal begins with the pixel values collected from the points selected on red channel image. Later CRT logic is applied on these values to obtain the desired crypto key. This generated key which is difficult to guess and secure, stronger, unique can be used for any symmetric encryption algorithm. This technique is helpful in obtaining variable length key as per the algorithm employed during encryption and decryption process.

Keywords Security · Key generation · Images · Image characteristics · Chinese remainder theorem (CRT)

K. Srinivas (✉) · N. S. Kumar · T. Sanathkumar · K. Rama Devi
Department of Computer Science and Engineering, Vaagdevi Engineering College, Telangana State, Warangal, India
e-mail: srinivas_k@vecw.edu.in

K. Rama Devi
Department of Computer Science and Engineering, Vaagdevi College of Engineering, Warangal, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_11

1 Introduction

Cryptography systems are often used in information technology to protect data that are sensitive vulnerable to unauthorized access during transmission. An algorithm which is a mathematical logic and the key as a parameter used by that logic are the main parts of cryptography. All cryptographic systems depend on the generation of secret, un-guessable and untraceable random numbers forming a key for cryptographic algorithms. At present, it is a challenging task for generating unpredictable keys which is forming an open wound in cryptographic systems. A new key generation procedure is built that can run on a wide range of applications meeting the requirement that the data and key by an adversary has a very low probability of guessing or determining.

In recent trends, communication advancements have led to the traveling of data in large volumes through electronic media. Therefore, development in cryptographic techniques and key architectures has been observed concentrating on the major issues of security. It was evident that the keys with required length are strong enough to overcome certain attacks like brute force etc. But still they are insecure as these keys are simple, easily predictable or obtained using few smart techniques. And also, insecure as they are shared among users, stored in insecure locations, using same password for various number of applications. This results to vulnerability of the system sharing confidential information to others.

Present, the work on information security turned toward merging images with a mathematical logic in cryptography as a means to improve overall security by eliminating the security issues. From a decade, images have become one of the commonly used entities in the field of information security. The success of its application in crypto security can indicate that many advantages could be gained. The key used during crypto process dependent on random bit generator. The generation process differs for each process like key from another key, from a passphrase, from a password or from a key agreement between two entities. Such keys used in any crypto algorithm may be vulnerable to an adversary to guess, predict or determine. Therefore, a key has to be produced which will be strong and resistant to such attacks. This document suggests for producing randomized key values from images and CRT making the key strong, secure and unique. And the length of each key depends on the algorithm used during encryption process. Number of keys assigned with different values can be generated from an image using the proposed key generation process.

The proposal, initially begins with conversion of RGB image to red channel image, then continues with application of CRT logic on the values obtained from red channel image which results into certain values that form a source for key selection randomly used for encryption algorithms.

2 Related Work

Our survey depicts the utilization of images in different applications like user authentication, encryption of images, image keys and generation of cryptographic key from images etc. The Paper [1], a novel authenticated framework based on image content is presented. The approach includes a hosting image that hides data without loss. Initially, the image is used for selection of features which are applied to digital signature algorithm to perform signing. The obtained authentication information from the signature and features is embedded in the selected image. Later by extracting the embedded information from the selected image, the original undisturbed image is obtained. Dhamija and Perrig [2] designed a graphical authentication technique using images. The paper mentions in its technique that the user has to prove his authenticity by identifying a set of pre-defined images. During registration, a set of images are randomly selected from the given image database. For successful login, the user has to authenticate by identifying the preselected images from the given set of images. This technique is based on the user memory capability. Monrose et al. [3] proposed a technique of generating a crypto key from password uttered by user. The user during authentication speaks the password. While uttering the password by the user, the voice is captured and verified to generate cryptographic key. This technique is bold enough to regenerate the key by allowing user repeating the password again. This technique may be typical when user utters the password with a change in voice.

Chang et al. [4, 5] proposed generation of crypto key from biometrics. Initially biometric features of user are extracted and collected. Then, these features are transformed to obtain a new feature of user dependent. Later using a stable mechanism of key generation, desired stable crypto key is generated by contributions from each new feature of authentic user. This technique may be complex if any of the biometrics gets disturbed. A new and secure scheme for integrating biometric-iris in crypto applications was proposed by Feng et al. [6]. The scheme follows generation of biometric error free key from repeated binary string and from iris codes obtained with two layers error correction technique. This technique is suitable for secure authentication of information. A methodology for generation of suitable length of crypto keys for 128-bit AES algorithm was proposed [5, 7]. First, a face image undergoes a number of iterations that results to set of bit sequences. Second, crypto keys are generated after applying entropy-based feature extraction coupling with Solomon error correction code on deterministic bit sequences. Implicit password-based authentication system (IPAS) for communication devices like mobiles etc., was proposed by Sadiq et al. [8]. This system comprises of two phases. During registration, the server extracts keywords from the personal information provided by the user. These keywords are represented as clickable areas on images which are stored in the server. The user during authentication has to click the right keyword on the image. The server provides the keyword and image that includes the keyword which are randomly selected and send to the user.

3 Methodology

It is a challenging task for generating unpredictable keys which forms an open wound in the design of cryptographic systems. A new key generation procedure is built that can run on a wide range of applications meeting the requirement that the data and key by an adversary has a very low probability of guessing or determining. At present a lot of research is going on to provide security for applications dealing with online information transmissions. One of the research outputs is using images in secure information transmissions. Numerous techniques were proposed for secure information transmission using images like steganography, encryption, authentication, etc. Our proposal is to use images during cryptography, that is, message encryption and decryption. How Images can be used to generate a variable length key used for encryption and decryption process.

3.1 Key Generation Technique

In cryptography, generation of cryptographic key for an algorithm has great importance. This section explains the application of CRT on the generated cryptographic key from an image which is used in encryption and decryption process. Initially, a group of images displayed where an image has to be selected and resized. This opted image undergoes a mechanism to produce a red channel image. Certain points are pointed on this red channel image and pixel values of corresponding pointed points are collected forming a matrix. Randomly a set of values are picked from matrix where the set is a predefined number. These randomly picked values undergo relative prime number verification. The resultant set is applied to CRT that results to an output which forms a source for key selection. The application of mathematical logic CRT, makes the key strong, unique and secure. Variable key length depends on the algorithm used during encryption and decryption process.

CRT: Chinese Remainder Theorem

- Let $r_1, r_2 \ldots r_n$ be relatively prime in pair and positive integers.
- Let $q_1, q_2 \ldots q_n$ be any values of integers. Then the linear congruence's system in one variable is given by

$$\begin{aligned} S &\equiv q_1 \bmod r_1 \\ S &\equiv q_2 \bmod r_2 \\ &\vdots \\ S &\equiv q_n \bmod r_n \end{aligned}$$

has a solution of modulo $r_1, r_2 \ldots r_n$ in unique.

Applications CRT includes:

- Used in the construction of sequence numbering of Godel's incompleteness theorems.
- Used for reduction computations in Fast Fourier transformations in Good Thomas algorithm.
- Used in RSA implementations of HTTP's signing certification.
- Interpolation in Lagrange, Polynomial multiplication etc. are few other applications.

3.2 Methodology

Step 1: Initially an image (Is) is to be selected from the displayed set of images. The image selected (Is) is resized which undergoes a mechanism in order to change into red channel image (Ir).

Step 2: Certain number of points are made on the red channel image (Ir). The pixel values of these pointed points are obtained and stored to form a matrix (F_r).

The points pointed on the Ir are limited to a number i.e. $M_p = 1, 2, \ldots k$. Our methodology takes $k = 8$ i.e. $M_p = 1, 2, \ldots 8$. Therefore, a matrix is formed with size ($M_p \times 3$) i.e.

$F_{r(Mp \times 3)}$ taking the pixel values of pointed points on Ir into consideration.

Step 3: T number of values are selected from F_r of which say $T1$, $T2$, $T3$ from T.

Step 4: $T1$, $T2$ and $T3$ undergoes for relative prime nature verification.

Step 5: These relative prime numbers are applied to CRT algorithm.

Step 6: Output of CRT is displaying S number of values. S—a source for selection of random variable length keys. The selected key from S is used for encryption/decryption process.

$$S = \text{CRT}(T1, T2, T3)$$

//$T1$, $T2$, $T3$ are selected from T

$$S = \{S1, S2, S3, S4 \ldots\}$$

//S output of CRT i.e., set of values $S1$, $S2$, $S3$, $S4$ … Sr.

Step 7: The size of S increases if T values increases thereby generating large number of strong, non-repeated values which can be used for encryption/decryption process.

Fig. 1 An selected Image (Is) which is resized to particular dimensions referred as resized Image. Later this resized image undergoes conversion process resulting to red channel Image (Ir). This red channel image is used for mouse pointed points by the end user as shown above

3.3 Experimental Results

Input: Select an image (Is) from displayed set of images and resize it.
Results:

1. Red channel image (Ir) from resized image (Fig. 1).
2. Formation of Matrix from pointed points on red channel image.

$$F_r = \begin{bmatrix} 162\ 0\ 0 \\ 245\ 0\ 0 \\ 223\ 0\ 0 \\ 202\ 0\ 0 \\ 255\ 0\ 0 \\ 85\ \ 0\ 0 \end{bmatrix}$$

T a variable that defines a number used for selection of values from $F_{r.}$

$$T = 162\ 245\ 223\ 202\ 255\ 85$$

3. These values undergo for relative prime nature verification.

$$\begin{aligned} &T1 = 152 \\ &T2 = 245 \\ &T3 = 223 \text{ are relative prime numbers} \end{aligned}$$

4. These relative prime numbers are applied to CRT.
5. Output of CRT is displaying S number of values. S—a source for selection of random variable length keys as shown.

$$\begin{array}{cccccc} T1(r1), & T2(r2), & T3(r3) & 162 & 245 & 223 \\ q1 & q2 & q3 & 2 & 3 & 5 \end{array}$$

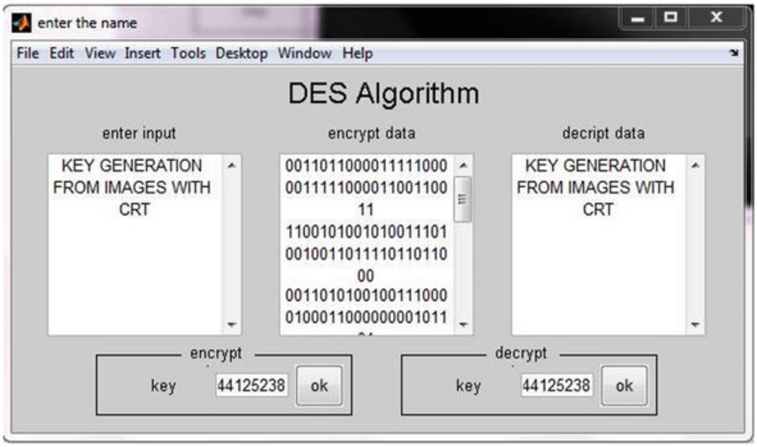


Fig. 2 Implementation of proposed method using DES algorithm. One of the symmetric Encryption/Decryption algorithms i.e. DES Algorithm is used for the implementation of our method by randomly selected key i.e. $S1 = 44{,}125{,}238$ from S as shown in Fig. 2. The key selected randomly forwarded for decryption process via trusted thirty party

$$S = \mathrm{CRT}\ (162,\ 245,\ 223)$$

$$S = \begin{bmatrix} 8721758 & 17572628 & 26423498 \\ 35274368 & 44125238 & 52976108 \\ 61826978 & 70677848 & 79528718 \\ 88379588 & 97486768 & 106582729 \end{bmatrix}$$

A selected number randomly from Step V is used as a key during encryption/ Encryption process in DES algorithm.

4 Conclusion and Future Scope

This work concentrates on how to generate a crypto key that is unpredictable by a cryptanalyst and how to make the key more strengthened. In this considering the image feature, that is, color, cryptographic key is generated with the application of CRT. Such generated key is used for any encryption algorithm. Our approach is limited to three input values to CRT but by increasing the input values leads to good results. This concept limits the adversary in prediction of the key as the key is created from images and CRT. If such work is carried out on both the sender and receiver side could bring better results in information security.

References

1. Zou D, Wah C, Yun Q (2003) A content-based Image Authentication system with lossless data hiding. In: International conference on multimedia and expo, ICME'03, vol 3, 6–9 July 2003. IEEE
2. Dhamija R, Perrig A (2000) DéjàVu: a user study using images for authentication. In: Proceedings of the 9th USENIX security symposium, 14–17 Aug 2000, Denver, Colorado, USA
3. Monrose F, Reiter MK, Li Q, Wetzel S (2001) Cryptographic key generation from voice. In: Proceedings of the 2001 IEEE symposium on security and privacy
4. Chang YJ, Zhung W, Chen T (2004) Biometrics-based cryptographic key generation. In: IEEE international conference on multimedia and expo (ICME), vol-3, 27–30 June 2004. IEEE
5. Srinivas K, Janaki V (2021) Symmetric key generation algorithm using image based chaos logistic maps. Int J Adv Intell Paradigms Indersci. Indexed in Scopus (Elsevier)-FREE, Standard age of Journal-13 Yrs. (since 2008), ISSN (online): 1755-0394, ISSN (print): 1755-0386. https://doi.org/10.1504/IJAIP.2021.10034870
6. Feng H, Ross A, John D (2006) Combining crypto with biometrics effectively. IEEE Trans Comput 55(9). IEEE Computer Society
7. Khalaf AD (2016) Fast image encryption based on random image key. Int J Comput Appl (0975–8887) 134(3)
8. Sadiq A, Prakash V, Naveen C (2013) A novel image-based implicit password authentication system (IPAS) for mobile and non-mobile devices. Math Comput Modell 58:108–116. www.elsevier.com/locate/patrec

Study of Transfer Learning Models for Disposition-Based Music Recommendation Systems

Y. Sri Lalitha, Y. Gayathri, Yamsani Nitish, Devarakonda Naveen, Bhagavathula Krishna Vamsi, and Mekala Nithish Kumar

Abstract The mood of a person can be extracted from facial features. These facial features can be obtained either by a static input image or by a live capture through the system's video input device (camera or webcam). Music is something that involves many aspects of human life and connects people over any differences. For a period, many developers and researchers are conducting studies to detect various moods of a person and classify them by building appropriate models. In this work, identify the mood of the user and play the music from the playlist, which the user has previously listed and played in different moods. It performs real-time mood recognition and develops a music playlist based on mood labels. In various studies, many different techniques exist to understand an individual's feelings, through facial dispositions (Jeevan Nagendra Kumar et al. in IEEE—5th international conference on communication and electronics systems (ICCES), pp 736–741, [1]), and in this work, we are studying Transfer Learning models, CNN, and its variants for a better understanding of emotions. The algorithms include Mobilenetv2, VGG-16, and Resnet 50. The models thus trained will help in making Android-ML integration easier. The main objective of the project is to develop an application capable of changing the user's mood by using music as a catalyst with the help of the user's facial features.

Keywords Mood detection · Face recognition · Music recommendation · CNN · MobileNetv2 · VGG-16 · Restnet50

Y. Sri Lalitha (✉) · Y. Gayathri
Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India
e-mail: srilalitham.y@gmail.com

Y. Nitish · D. Naveen · B. K. Vamsi · M. N. Kumar
Department of IT, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_12

1 Introduction

Music has a therapeutic effect on human brains proven in many studies. It stimulates the portion of the brain that creates the dopamine hormone that influences a person's mood to be pleasant or feel good. It can change the emotional conduct of persons in a bad mood, or bad health. Music contains a behavioral furthermore as a neurological impact. That's to mention, music influences not just our mood but also our inability to manage our emotions. The emotion of a human can be identified from various actions. Like emotions can be recognized using voice, facial expression, by heart rate, etc. According to further investigation, emotion recognition using face is a bit easier. The emotion [2] of a human is displayed by the face and can be felt by the brain, which can be captured using electroencephalography (EEG) a recording of brain activity, from video into an image form. The recognition of the mood from image form can be done with efficient Classification Models. Deep learning models have shown high accuracy in image classification [3, 4]. Transfer learning is a breakthrough in Deep Learning models. This work-study Transfers learning models for Mood Detection.

2 Related Work

In the literature, various studies are proposed for image detection, and emotion identification [5]. Music has a therapeutic effect on people in bad moods, and bad health. This has been proven scientifically. Mood-based music playlist generation is a widely studied area. The image and emotion-detecting techniques add to this research work. Here is a brief list of such studies. Algorithms Viola-Jones and Haar cascade are used to detect the face. Deep Learning algorithms CNN, VGG 16, MobileNet V2, and Resnet50 are the classification models to validate face detection performance with good accuracy. A system in [6] build models on "CK (Cohn-Kanade)" and "CK+ (Extended Cohn-Kanade)" datasets, using Computer Vision Library "OpenCV" whose accuracy is approximately 83%. However, other systems have an accuracy of 88% trained on the dataset "KDEF (Karolinska Directed Emotional Face)" and the "VGG (Visual Geometry Group) 16" algorithm combined with the "CNN (Convolution Neural Network)" algorithm. The original "CK" dataset is viewed as a very beneficial inclusion to the current corpus due to its popularity and ease of access. It is also appealed on the bases of a completely autonomous system to be reliable for every human facial expression in a variety of real-world settings, most of the information is still needed. Therefore, it would need collaboration among all research labs to collect a large, consistent dataset across a wide variety of visual disparities to be collected and organized efficiently (which includes samples ranging from 5 to 10,000 on every action).

Nitisha et al. [3] performed an experiment based on the cross-database raw features. It was observed that the two datasets, the Mobile pictures dataset and Logistic Regression for testing RaFD (Radboud Faces) database, performed better when using raw features. By using the CK+ dataset [7] as the training set the accuracy achieved is 66% and 36% respectively. The accuracy of the SVM (Support Vector Machine) decreased once the additional characteristics (distance and area) were included. The approach performed better than SVM and several other algorithms when it was implemented for the training and testing sets. The average accuracy was 86% for the RaFD database and 87% for the CK+ database for a cross-validation of five times. The main emphasis was on the dataset's machine algorithm analysis and feature extraction. However, precise face-detection algorithms become essential when there are multiple people in the image. One of the works [8] was put to the test by drawing expression from either a previously saved image or a live feed provided by the system's camera. It was put into practice using Python 2.7, OpenCV, and NumPy. The objective was to develop a system that can evaluate a photograph and predict an individual's facial expression. The investigation showed that this procedure is workable and produces trustworthy findings.

The Music Recommendation System has been the subject of study as well [7, 9–13]. One such study [9] describes a basic method for classifying the mood of Hindi music that makes use of easy-to-extract audio elements. Using the tenfold cross-validation, the MIREX (Music Information Retrieval Evaluation eXchange) mood taxonomy provided an average accuracy of 51.56%. Additionally, an article [10] claims that the present music recommendation study is the outcome of the description of musical resources. It is asserted that present research suffers from a lack of systematic study of user behavior and demands, a low degree of feature extraction, and a single assessment metric. The situation was shown to be a key component in the customized music suggestion system. The accuracy of the recommendation findings was found to be significantly lowered when the weights assigned to each contextual component were equal.

According to a different study [11], their hybrid recommendation system method idea will succeed after their model has been sufficiently trained to detect the labels. Brain waves and audio characteristics from music are used by the technology to automatically extract user preference data and maintain user preferences in the personalized music recommendation service. Their work employs a very brief feature performance hit and a distance measure learning technique to address music genre categorization challenges. The suggested user preference classifier has an overall accuracy of 81.07% in the binary preference classification for the KETI AFA2000 music corpus. It was clear that the customer was happy when brainwaves were used. The vector is produced using low-dimensional projection and already existing audio attributes. The feature vector's dimensionality was reduced with only a slight difference.

2.1 Existing Systems

EMO Player: An innovative method called Emo Player (an emotion-based music player) helps the user in coping with his mood with the help of automatic music played according to the emotion of the user [2]. *SoundTree*: A music suggestion tool called Sound Tree may be integrated into another internet program. It makes advantage of people-to-people linkage based on the user's previous actions, such as having listened to or downloaded music [5]. *lucyd*: A music recommendation engine called lucyd was created by four students who are graduates of UC Berkeley in the master program—Master of Data Science (MIDS). Users may use any phrases they like to search for music suggestions on lucyd [6]. *Reel* Time.AI: The user must subscribe for this system to function. After that, the user may upload pictures of major gatherings like malls, theatres, and restaurants. The algorithm then distinguishes between joyful and sad moods. It determines which faces conveying joy and which ones convey sadness, and it interprets the scenario based on the expressions on the faces of everyone in the room. *Music.AI*: It proposes music based on the user's selected mood using the list of moods as input. It combines content-based and collaborative filtering approaches for filtering information. When recommending music, factors including emotion, timing, atmosphere, and learning history are taken into consideration [8]. The argument is whether can we find the best of the existing models and come up with an efficient solution.

3 Proposed System Architecture

This chapter proposes a comparative study of the algorithms which are used for building the model for emotion recognition and classification. The study includes the comparison between the CNN algorithms MobileNetv2 [12], VGG 16, and Resnet50. The algorithms are trained and tested on the same datasets, the accuracy is compared, and the best algorithm is used for the development of the application. The application is built using the pygame module. The Haar cascade algorithm is used for face detection. The application uses the camera of the system and detects the face.

3.1 Architecture

The system architecture is made up of three parts:

Face detection: The face of the user is detected by taking a snapshot of the user from the live feed through the system camera.

Mood Detection: The mood will be detected from the mood/class labels.

Music recommendation: After the mood is detected the music is recommended according to the user playlist which is stored in the database.

Figure 1 describes the overall process, methods, and functions from input to output of the proposed system. The snap is taken for mood detection. The mood is detected by using a trained model of type, h5 file. According to the mood label, the music

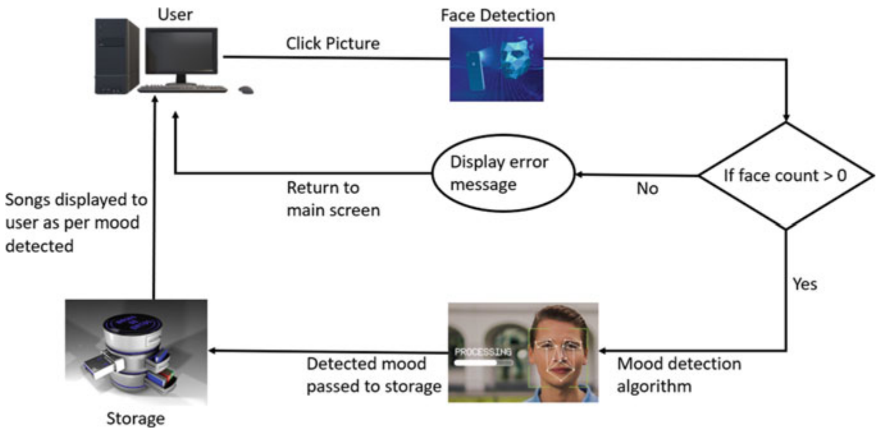


Fig. 1 Proposed system architecture

playlist will be played from the database. The user can create or save their choice of music playlist in the database.

4 Implementation

4.1 Dataset Collection

Dataset is collected from the Kaggle repository and through GitHub. FER 2013 [4] is the dataset, which contains greyscale images of size 48 × 48 pixels. The dataset obtained from GitHub is an emotion recognition dataset [1] which are in different resolution. Both datasets are approximately 60 MB each.

4.2 Preprocessing

In the process of machine learning, preprocessing is essential. Preprocessing data lead to improved outcomes. To remove noisy photos from the dataset, various text preprocessing methods are investigated. From a collection of huge datasets, a variety of image-processing techniques are used to extract clear and noise-free pictures [8.1][8.2]. We have applied Data augmentation and image resizing [8.3].

Data Augmentation: To solve the problem of imbalanced data, Data Augmentation steps were applied to the whole dataset (Zooming, Rescaling, Flipping, etc.).

Resizing the image: To ensure consistency among all images, each image was resized to (224,224).

4.3 Model Construction

The following three models were created using the transfer learning methodology: VGG16, ResNet50, and MobileNetV2. Transfer learning is the ability to learn from experiences and apply the acquired knowledge in new environments [14]. The model was built using this process because TL offers benefits like computation efficiency, optimal resource utilization, increased effectiveness, etc.

VGG16: [13] The 2014 ILSVR (ImageNet) contest was won by a convolutional neural network. To categorize 1000 distinct classes, it was trained on 1.2 million photos. Figure 2 shows the VGG16 architecture. It has around 138 million characteristics and is 16 layers deep. It is frequently employed in categorization issues and is simple to apply to utilize transfer learning. It is publicly accessible in the Keras library and includes weights that have previously undergone training using the Image dataset. This significantly cuts down on the training phase. In this work, a pre-trained VGG16 model with 138 million trainable parameters was imported from the Keras library. To fine-tune, all the layers were made non-trainable, sequential models and eliminated the last layer to shorten training time and output classes. With just 23 outputs, we inserted the dense layer as the last layer. The trainable parameters were dramatically reduced from 1.34 million parameters to 94,231 parameters.

MobileNetV2: [12] Is a convolutional neural network, that claims to be responsive to mobile devices. It is based on residual connections between bottleneck levels and an inverted residual structure. As a source of nonlinearity for the intermediate expansion layer filters, light depth-wise convolutions are present. The architecture of MobileNetV2 is shown in Fig. 3. It contains over 3.4 million trainable parameters and is 53 layers deep. A MobileNet V2 pretrained model was imported into this project from the Keras library. We made all the layers non-trainable, made the model sequential, and eliminated the last layer to shorten training time and output classes. With just 23 outputs, we inserted the dense layer as the last layer. This technique for fine-tuning greatly assisted in lowering the trainable parameters.

ResNet50: [15] A convolutional neural network called ResNet-50 (Residual Network) provides the foundation for several computer vision applications, including object identification and image segmentation. Below is a diagrammatic depiction of ResNet50. It contains 23,587,712 parameters and 50 layers. Additionally, it was named the victor of the 2015 Image Net Challenge. It also addressed the vanishing

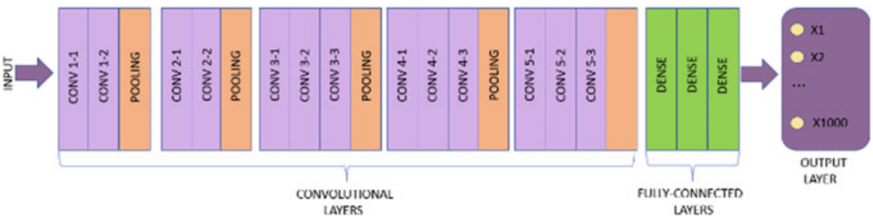


Fig. 2 VGG 16 architecture

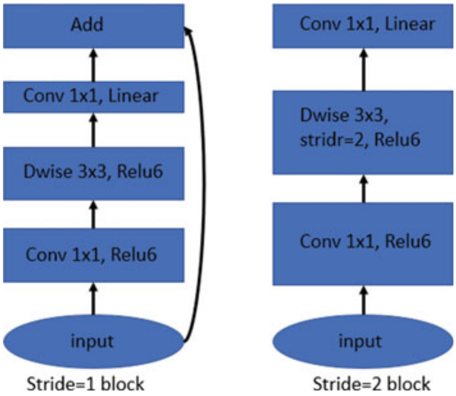


Fig. 3 MobileNet V2 architecture

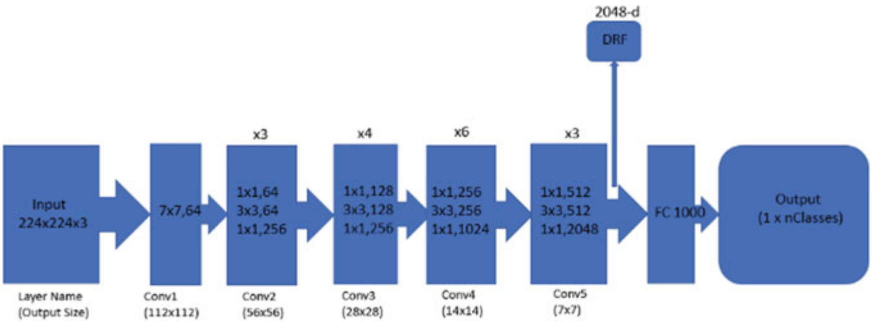


Fig. 4 ResNet50 architecture

gradient problem, which was a common one in deep neural networks. We used the ResNet 50 pre-trained model that was loaded from the Keras library for this project. We made all the layers untrainable, made the model sequential, and eliminated the last layer to decrease the training, the time, and the output classes. With just 23 outputs, we inserted the dense layer as the last layer. The trainable parameters were dramatically reduced by using this approach of fine-tuning, going from 23,869,463 to 2817 (Fig. 4).

4.4 Training Model

To extract the best feature 70% of the dataset is utilized for the creation of a training dataset and models are trained using this dataset. The training set contains more than 22,968 images for training, and 5741 images are used for validation. 100 epochs, batch_size as 32 with Early Stopping mechanism is applied on each developed model. The model .fit () function will be used for training the created model (Table 1).

Table 1 Training performance of three models

Model	Validation accuracy (%)	Validation loss
VGG16	81.56	0.46
MobileNet V2	84.45	0.31
ResNet50	83.43	0.42

Table 2 Test performance of three models

Algorithm	Test accuracy (%)
VGG16	85.24
MobileNet V2	87.67
ResNet50	86.82

Testing Model: After model training. On the test dataset, which made up 30% of the total dataset, each model was assessed. This stage determines if the model can effectively learn from and categorize the data. The observation is depicted in Table 2. MobilenetV2 depicted high accuracy, with 94% approximately.

5 Experimental Results

Accurately identifying human emotions or moods is difficult because each person has distinctive facial features. However, it can be partially identified with the right facial expressions. The instrument should have an excellent resolution to facilitate detection. Upon the selection of the detect button, the model effectively "Detects your mood" the detection platform will appear, and the snap of the face is taken by the camera of the device and takes the snap as the input for mood detection. Figure 5 is the home page of the developed application.

Once the input is received the mood will be detected accordingly. The system identification of different moods is depicted in the below figures. Figure 6 shows the "Neutral" mood, "Angry Mood" is depicted in Fig. 7, "Surprise Disposition" is identified in Fig. 8, and "Happy Mood" can be observed in Fig. 9.

According to the detected mood, appropriate music is recommended automatically. The system displays the playlist on the screen, based on the user's previous playlist interests and the kind of music acoustics the system categorizes the playlist with different emotions.

Once the system identifies the disposition of the person, Fig. 10 will be displayed with appropriate emotion, and confirms the mood of the person on the interface with the proposed "Play List". Displaying the correct "Playlist" is the outcome of this work.

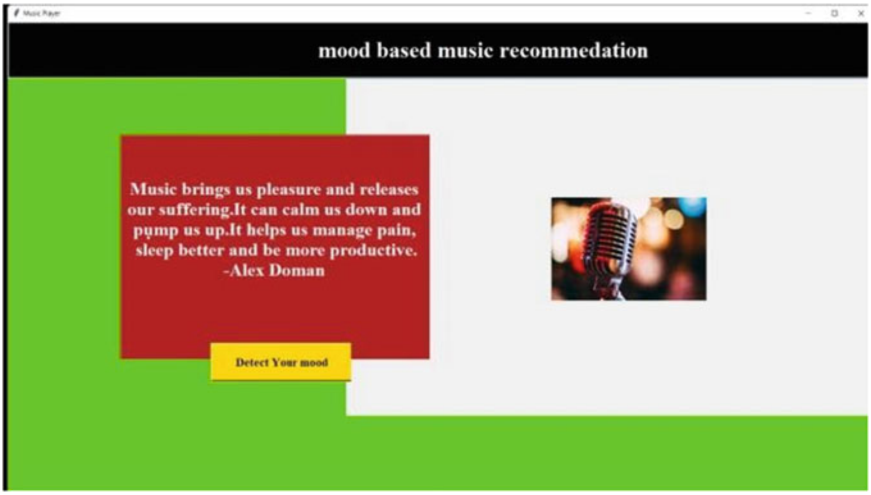


Fig. 5 Application interface

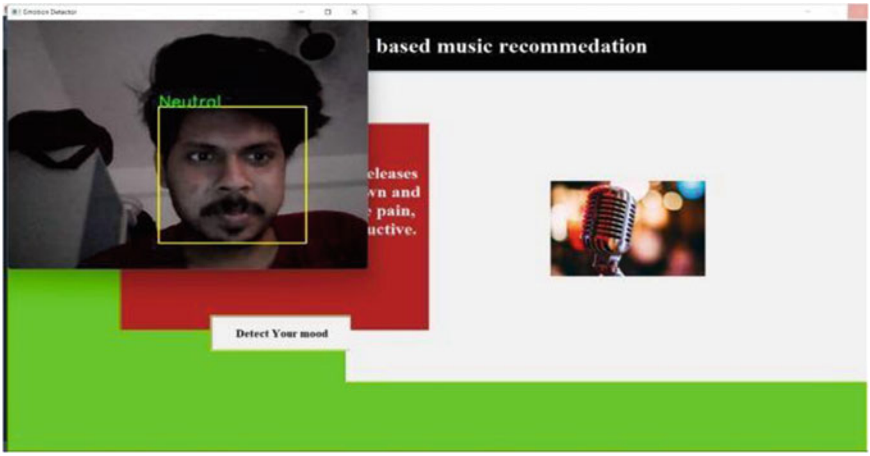


Fig. 6 Disposition: "Neutral"

5.1 Graphs

The training and validation accuracy obtained by the VGG16 model is visualized in Figs. 11 and 12 depict the loss of the VGG 16 Model, the *x*-axis displays the number of epochs, and the *y*-axis represents the accuracy/loss values.

Figure 13 shows the ResNet50 model's training and validation accuracy, with the *y*-axis representing accuracy values and the *x*-axis representing the number of epochs.

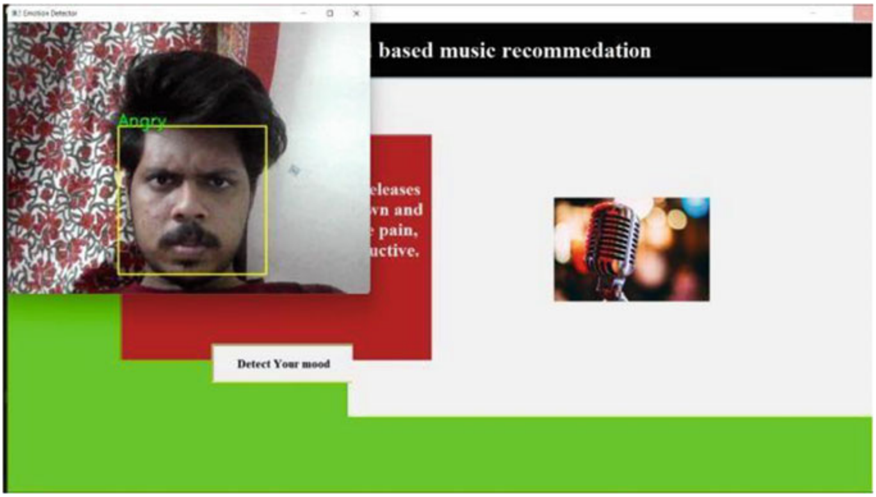


Fig. 7 Disposition: “Angry”

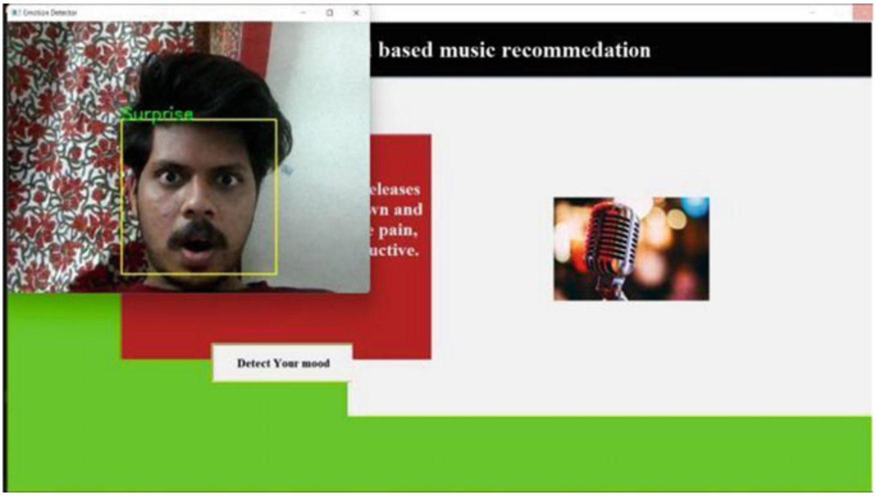


Fig. 8 Disposition: “Surprise”

Figure 14 shows the training and validation loss of the ResNet50 model, with the number of epochs on the x-axis and the loss values on the y-axis.

Figure 15 visualizes the training accuracy and validation accuracy of the MobileNetV2 model, where the x-axis resembles the epochs trained and the y-axis resembles the accuracy values.

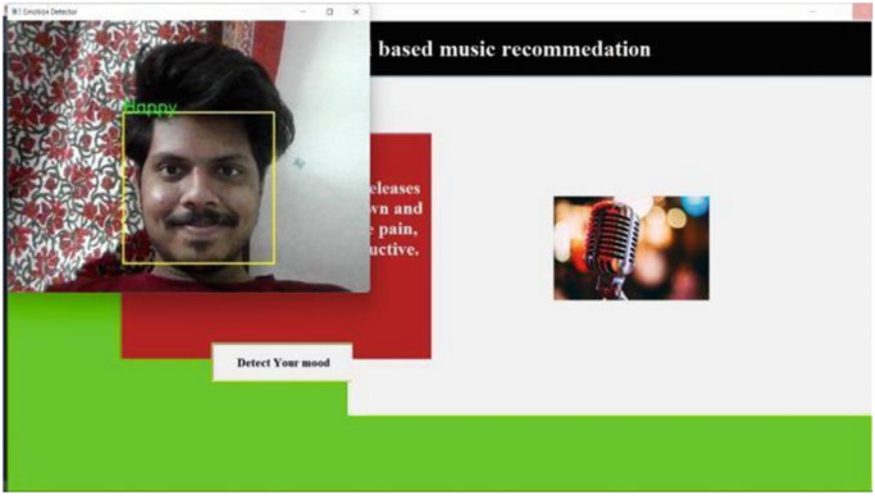


Fig. 9 Disposition: "Happy"

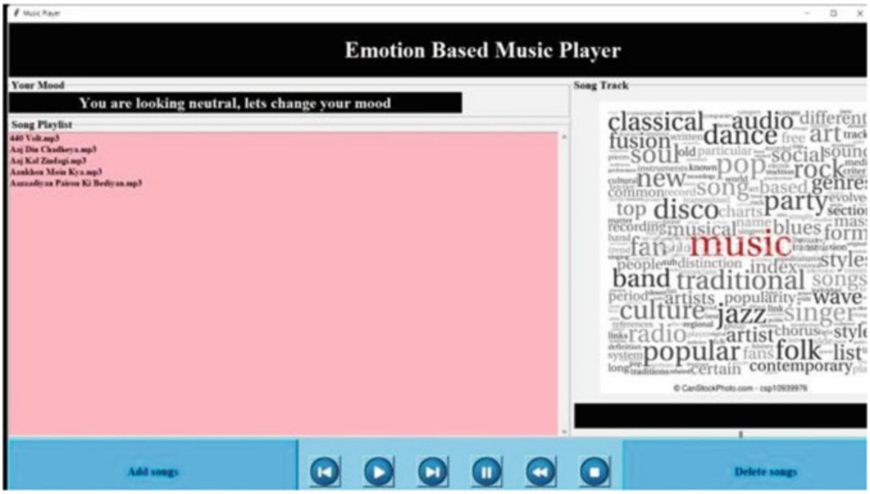


Fig. 10 Application playlist

The training and validation loss of the MobileNetV2 model is shown in Fig. 16, where the *x*-axis represents the number of epochs, and the *y*-axis represents the loss values.

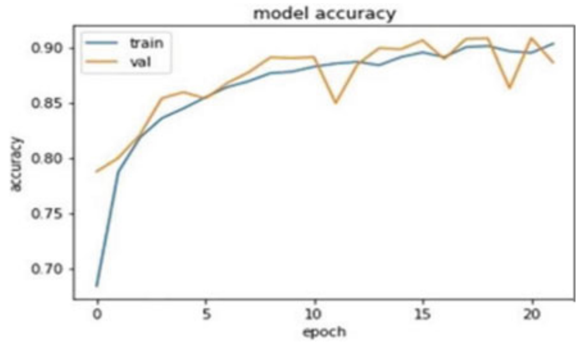


Fig. 11 VGG 16 model accuracy

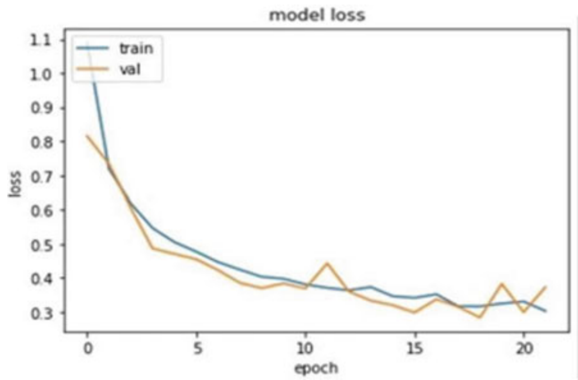


Fig. 12 VGG 16 model loss

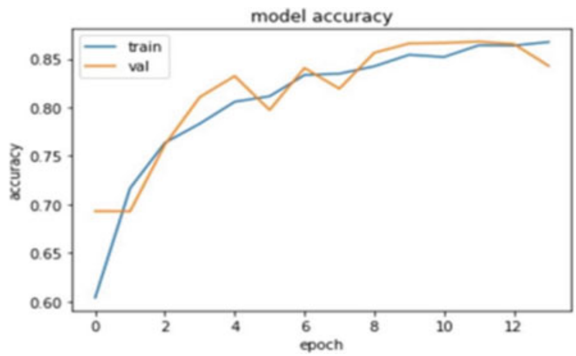


Fig. 13 ResNet50 model accuracy

6 Conclusions and Future Scope

Our system is an application that helps in changing the mood of a user. Our model is based on extractive methods. The access time for mood detection and face detection has reduced latency. The datasets we used in the building of the

Fig. 14 Resnet model loss

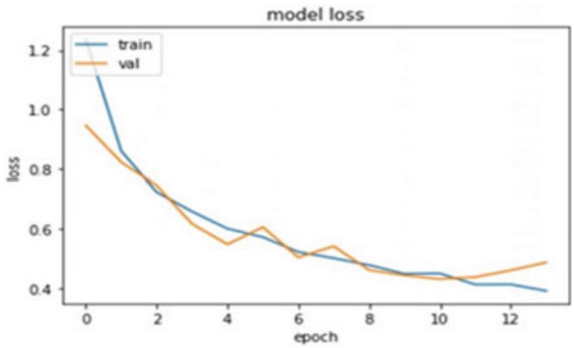


Fig. 15 MobileNetV2 model accuracy

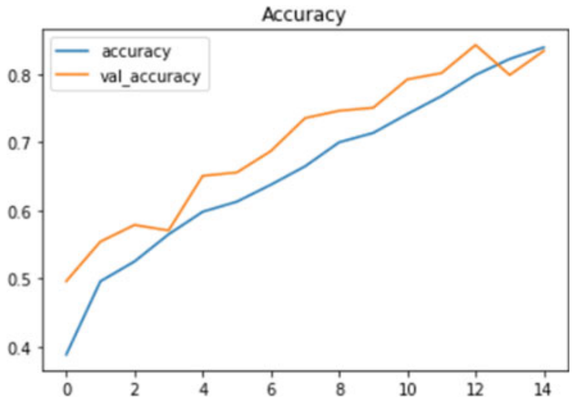
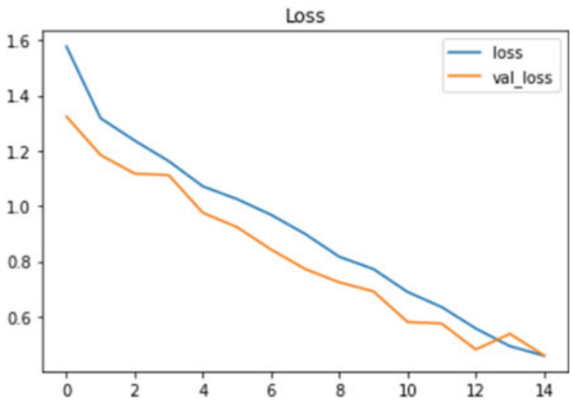


Fig. 16 MobileNetV2 model loss



models are first trained using the CNN base model, with a training accuracy of 78%, it took 4.5 min in for building the model and 4 min to detect the emotion for test data. This work is a comparative study of Variants of the CNN model (“VGG16, ResNet50, and MobileNetV2”) for emotion detection. It is observed that “MobilenetV2” has depicted 87% accuracy. The MobilenetV2 model is used for developing the application for music recommendation players with the user’s emotions.

The difficulty of this work is to create an AI-based database as per user’s interests. Music has a healing effect, and mood change behavior, and the appropriate recommendation of music, in native languages will soon become an important mobile application down the line. Hence many deep learning and AI-based models should be developed especially in language-specific recommendations. In our future work, we will explore multi-language music recommendations.

References

1. Jeevan Nagendra Kumar Y et al, Supervised machine learning approach for crop prediction in agriculture sector. In: IEEE—5th international conference on communication and electronics systems (ICCES), pp 736–741. ISBN: 978-1-7281-5370-4
2. Hemanth P et al (2018) EMO PLAYER: emotion based music player. Int Res J Eng Technol (IRJET) 5(4):4822–87
3. Nitisha et al (2018) Facial emotion recognition using machine learning. Master’s projects, 632
4. Sri Lalitha Y et al (2021) Neural network modelling of speech emotion recognition system. In: 2021 in E3S web of conference, ICMED. https://doi.org/10.1051/e3sconf/202130901139
5. Naik DA et al (2022) Analysing histopathological images by using machine learning techniques. Appl Nano Sci
6. Abdul A et al (2018) An emotion aware personalized music recommendation system using a convolutional neural networks approach. Appl Sci 8(7):1103
7. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression. In: 2010 IEEE computer society conference on computer vision and pattern recognition—workshops, San Francisco, CA, USA, pp 94–101
8. Hussain SA, Al Balushi ASA (2020) A real-time face emotion classification and recognition using deep learning model. J Phys Conf Ser 1432 012087
9. Patra BG et al (2013) Automatic music mood classification of Hindi songs. In: Proceedings of the 3rd workshop on sentiment analysis where AI meets psychology (SAAIP 2013), IJCNLP 2013, pp 24–28
10. Lee J et al (2018) Music recommendation system based on genre distance and user preference classification. J Theor Appl Inf Technol 1285–1292
11. Kaufman Jaime C (2014) A hybrid approach to music recommendation: exploiting collaborative music tags and acoustic features, UNF Digital Commons, University of North Florida
12. Sandler M et al (2018) MobileNetV2: inverted residuals and linear bottlenecks. In: IEEE conference on computer vision and pattern recognition, pp 4510–4520
13. Takur R (2019) Step by step VGG16 implementation in Keras for beginners. Towards Data Science 2019
14. Ramasuri Appalanaidu CH et al (2021) Music recommendation system with advanced classification. Int J Eng Res Technol (IJERT) 10
15. Dwivedi P (2019) Understanding and coding a ResNet in Keras. Towards Data Science 2019

Detecting Anomalies in the Virtual Machine Using Machine Learning Techniques

H. Priyanka, Manjula Gururaj Rao, and M. S. Apoorva

Abstract There is an increase in demand for cloud computing services, and optimizing the resources is crucial. Conventional algorithms are not adequate to address issues with the cloud, including uneven task distribution in virtual machines or inefficient resource delivery to cloud customers. To significantly enhance the outcome of cloud apps and address the aforementioned difficulties, it is necessary to investigate further sophisticated ways. This study looks into the most recent machine learning techniques that can handle the difficulties in a cloud context. A paradigm monitoring and identification of errors in Virtual Machine (VM) resource utilization is put forth in this study. The suggested methodology called IF-SVM (Isolation Forest-Support Vector Machine) is able to identify the errors. The VM workload trace from PlanetLab is used to test and train the samples. During this process, the concepts of ML techniques such as SVM and IF and the VM resource matrix is used. F1-score of 0.87 is able to achieve by using the SVM for the time series of 1 h and 0.91 F1-score for Isolation Forest for the 50 VMs. This result demonstrates the effectiveness of both methods for the model, but SVM outperforms well in terms of classification in comparison with Isolation Forest.

Keywords Cloud computing · Virtual machine (VM) · Support vector machine (SVM) · Isolation forest (IF)

1 Introduction

Information storage and accessing are one of the prominent fields in the research. Many methods have established for accessing and storing the information. One of them is cloud computing. In effort to provide speedier innovation, flexible resources,

H. Priyanka (✉) · M. S. Apoorva
PES University, Bengaluru, India
e-mail: priyankahsachin@gmail.com

M. G. Rao
NITTE (Deemed to Be University), NMAM Institute of Technology Nitte, Karkala, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_13

and scale economies, cloud computing is the distribution of computer services via the Internet ("the cloud"), including analytics, servers, software, storage, networking, databases and knowledge.

The process that moves the virtual computers between data centers is known as virtual migration. Multitenancy is a process of dividing the resources of the server to various users according to their need.

Because of the public cloud's multi-tenancy, attackers leverage the migration module's capability to obtain the resources of the VM [1, 2]. The aim of this chapter is to secure the cloud architecture by supplementing the current security measures for the cloud. Especially for IAAS (Infrastructure-As-A-Service) [3] a cloud platform concept where services can be readily affected due to its multitenancy, which exposes users to more security concerns.

In order to safeguard the VM resources, this study explores the usage of "Machine Learning (ML)" approaches in the identification of irregularities in VM resource utilization within the cloud. Particularly in a public cloud architecture with multitenancy [4], there may be vulnerability exploitation and a risk of compromise with VM-to-hypervisor or VM-to-VM communication. Tackle the problem, research suggests using unsupervised ML algorithms to create a model that will track and identify irregularities in the utilization of VM resource.

Following is how the remaining sections of the paper are arranged: Discussions of related works, the suggested techniques are shown in Sects. 2 and 3. The results and the discussion is done in the Sect. 4. In Sect. 5, the conclusion and suggestions for additional research are covered.

2 Related Work

The majority of network architectures and methodologies, however, are concentrated on protecting the network's security from assaults (such denial-of-service attacks), failures, and system performance detections [5]. The authors also suggest ML and statistical techniques to improve intrusion detection systems [6, 7]. The virtualization technology and VMs in the cloud, and authors in [8] have addressed how security vulnerabilities involving resources of the VM can be identified using ML and statistical models, but the models may use a single metric such as utilization of CPU, making it difficult to identify malicious attacks on the metrics of the VM.

The architecture of cloud computing must include security feature such as intrusion detection system (IDS), which acts as a protection against "intruders and attackers". Mehmood et al. [9] performed an examination of different IDS used in cloud as a precaution for security and their shortcomings. Although IDS is a crucial protection measure, cloud service providers mostly employ them to find an attacks at network-level [8].

The VM metrics such as throughput for reading disc and throughput for disc writing are merged and used in the model-building process. As also noted in Tiwari et al. in [8], this can identify hidden irregularities that a single parameter could overlook.

A novel hybrid model has been proposed by the authors Junaid et al. [7] to provide a proper load balancing solution in cloud computing, with comprehensive data. It uses file type formatting to identify the number of files available on the cloud. Considering different file formats such as audio, video, text maps, and cloud images, the classification is carried out using the Support Vector Machine (SVM). The drawbacks of proposed work are: it can be enhanced by using deep learning methods and other swarm-based techniques to solve a load balancing problem. In this study, we present a method for detecting irregularities in the utilization of VMs that combines measurements from various virtual machines as input to a ML algorithm.

3 Proposed Technique Using Machine Learning Approaches

The system must be continuously monitored and investigated for deviations from the typical pattern of VM resource usage, even if a variation in resource usage patterns may not always signal an attack because authorized customers may still be utilizing their virtual machines for legitimate workloads. Appropriate decisions need to be taken if irregularities occurs in VM's resources usage.

The monitoring of VM metrics is essential, as it gathers and extract the necessary metrics such as CPU utilization, throughput, and memory consumption to determine the behavior of resource utilization that will serve as the foundation for the ML model. It is necessary to collect and save historical usage information for virtual machines in order to evaluate it and identify the design. After gathering the data for qualitative analysis, extraction of features is done to make sure that all of the metrics fall within the same range of values, thus protecting the model's effectiveness and classification accuracy. Later in the evaluation step, it will find the anomalies where there is any change in VM resource pattern (Fig. 1).

(a) Collection of Data

The PlanetLab datasets [2, 10, 11] are CoMon initiative, and are collected only for 10 days from 500 distinct places around the world. So, in line with that to carry

Fig. 1 Proposed architecture

out the research activity ten days data are used. The data on CPU utilization was collected from 500 distinct places around the world which are running on the server from thousands of VMs. Therefore, every trace has a data of 24 h * 60 min and the interval of CPU usage is 5 min which is equal to $(24 * 60)/5 = 288$ entries.

Once the features had been scaled, these measurements were resampled into hourly series, and employed for training the model and also testing of the model.

The dataset for testing may possess the identical properties and it's derived from the exact workload trace. The proposed model was tested using a combination of test data and regular data points of the dataset.

It is utilized to evaluate the suggested model along with the test data. A portion of the data points have a fluctuation in their range of values (which serve as the anomalies) that differs from those in the typical (training) datasets after the preprocessing is finished to test the model.

(b) Implementation

The hybrid model was created using the two unsupervised machine learning techniques such as "Isolation Forest and Support Vector Machine" known as IFSVM. The Scikit-Learn [12, 13] module of python contains both the algorithms. The two machine learning algorithms that were employed are described briefly below:

(i) Isolation Forest

A data point that considerably varies from several other data points in the dataset is referred to be an outlier. Identifying outliers in the data—that is, points that are substantially varies from the bulk of the other points—is the method of anomaly identification. In unsupervised learning model, it searches the data for anomalies by focusing on outliers. Isolation Forest [14] is built on the Decision Tree algorithm. In the initial stage, it selects a feature at random from the available features, and then it distinguishes outliers by dividing the value between the minimum and maximum values of that feature. Data from the 1st stage are passed through separation trees to calculate the incongruity score in the testing phase also known as the 2nd phase.

(ii) SVM

During training, SVM [15] uses a hyper-plane to define a limit for data points. Iteratively, the limit is set using an optimal hyper plane. When testing, it determines if fresh data points fall inside the training boundary and labels them as either normal (1) abnormal (2).

Input = Extracted features of virtual machine

Output = classification [normal, abnormal]

//CPU - CPU usage

1. Initialization
2. Read input Data
3. If input Data ["feature"].data type! = right_data_type

Then: input Data ["feature"].datatype inputData ['feature].right_data_type end if

4. ExFeature inputData ["timestamp","CPU"]
5. ExFeature.Resample_to_time_series () //feature scaling of the individual metrics of the extracted features is done in 6-9
6. ExFeature ["CPU"] FeatureScale ("CPU")
7. Model SVM () // model IsolationForest ()
8. Model. Train (exFeature)
9. Outcome model. Test (exFeature)
10. Output outcome. Confusion matrix ()
11. Output outcome. Precision ()
12. Output outcome. Recall ()
13. Output outcome.F1-Score ()
14. End

4 Results and Discussion

Obtaining the raw data in the proper format is crucial for the evaluation. For this reason, data exploration throughout the preprocessing phase was essential. For the algorithm's convergence and to confirm the accuracy, the data features were scaled (normalized) to remain in the same range. As can be observed from the data visualizations in Fig. 2a, b, the metrics of normalization has no effect on or alters the pattern of the data points from the data which is original. The min–max scalar (normalization), which can be seen in the method in step 6 and is employed in steps 6 to 9 of the algorithm, is used for feature scaling.

Precision: This demonstrates how well the system foresees the abnormal events. It is written as:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

Recall: Another name for recall is sensitivity. This demonstrates how well the algorithm forecasts (positive) typical occurrences. As shown below:

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

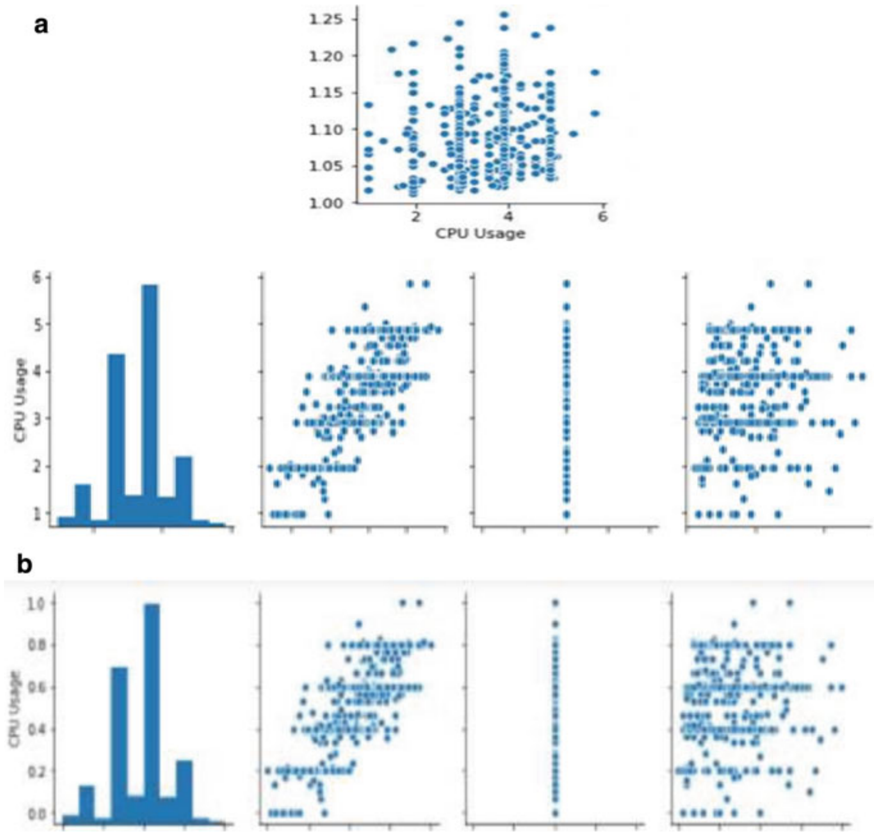


Fig. 2 **a** Unnormalized data. **b** Normalized data

The weighted average of memory and precision is the F1 score. It is written as (Table 1):

$$F1\text{ - score} = 2 * \frac{\text{Precision} * \text{recall}}{\text{Precision} + \text{recall}}$$

Table 1 Table of computation

	VMs count	Precision	Recall	F1-score
SVM	50	1.0	0.80	0.91
Isolation forest	50	1.0	0.81	0.87

5 Conclusion

Due to the common infrastructure that the cloud computing concept is based on its remote internet access, it is vulnerable to numerous internal and external risks and attacks. In order to minimize any potential assault that could adversely affect the cloud's ongoing operations. The proposed model uses metrics like CPU for VM resources to quickly identify and isolate resource usage irregularities. Additionally, it enhances or supplements the security measures that is already placed in the cloud. The results of the execution demonstrated that the proposed strategy has a greater rate of success with IF-SVM. The SVM is more efficient than IF, which is observed from the obtained results.

References

1. Boutaba R, Zhang Q, Zhani MF (2014) Virtual machine migration in cloud computing environments: benefits, challenges, and approaches. In: Communication infrastructures for cloud computing: IGI global, pp 383–408
2. Priyanka H, Cherian M (2020) Efficient utilization of resources of virtual machines through monitoring the cloud data center. In: Bindhu V, Chen J, Tavares J (eds) International conference on communication, computing and electronics systems. Lecture notes in electrical engineering, vol 637. Springer, Singapore
3. Priyanka H, Cherian M (2022) Energy-efficient resource management of virtual machine in cloud infrastructure. In: Buyya R, Garg L, Fortino G, Misra S (eds) New frontiers in IoT and cloud computing for post-COVID-19 era. Springer Nature, Switzerland. ISBN: 978-3-031-05527-0. https://doi.org/10.1007/978-3-031-05528-7_4
4. Prashanthi M (2016) Analysis of security issues in virtualization cloud computing. Int J Comput Sci Mob Comput IJCSMC 5(8):274–281
5. Nguyen H, Tan Y, Gu X (2011) Pal: Propagationaware anomaly localization for cloud hosted distributed applications. In: Managing large-scale systems via the analysis of system logs and the application of machine learning techniques, pp 1–8
6. Modi CN, Patel DR, Patel A, Muttukrishnan R (2012) Bayesian classifier and snort based network intrusion detection system in cloud computing. In: 2012 third international conference on computing, communication and networking technologies (ICCCNT'12). IEEE, pp 1–7
7. Junaid M, Sohail A, Rais RNB, Ahmed A, Khalid O, Khan IA, Hussain SS, Ejaz N (2020) Modeling an Optimized Approach for Load Balancing in Cloud. IEEE Access 8:173208–173226
8. Tiwari T, Turk A, Oprea A, Olcoz K, Coskun AK (2017) User-profile-based analytics for detecting cloud security breaches. In: 2017 IEEE international conference on big data (big data). IEEE, pp 4529–4535
9. Mehmood Y, Shibli MA, Habiba U, Masood R (2013) Intrusion detection system in cloud computing: Challenges and opportunities. In: 2013 2nd national conference on information assurance (NCIA). IEEE, pp 59–66
10. Priyanka H, Cherian M (2021) Effective utilization of resources through optimal allocation and opportunistic migration of virtual machines in cloud environment. Int J Cloud Appl Comput (IJCAC) 11(3):72–91. https://doi.org/10.4018/IJCAC.2021070105. (Web of Science, SCI indexed)
11. Priyanka H, Cherian M (2022) Metaheuristic based resource management approach towards minimization of energy consumption in cloud data centres. J Optoelectron Laser 41.7:676–687

12. ScikitLearn (23-03-2021) Novelty and outlier detection. Available: https://scikitlearn.org/stable/modules/outlier_detection.html
13. ScikitLearn. One class SVM. Available: https://scikitlearn.org/stable/modules/generated/sklearn.svm.OneClassSVM.html
14. Liu FT, Ting KM, Zhou Z-H (2012) Isolation based anomaly detection. ACM Trans Knowl Discov Data 6(1):1–39
15. Chang C-C, Lin C-J (2001) LIBSVM: a library for support vector machines. ACM Trans Intell Syst 2(3):1–27; Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RCJNc (2001) Estimating the support of a high-dimensional distribution 13(7):1443–1471

Human Facial Image Generation from Textual Descriptions Using StyleGAN

E. Eswar Reddy, Mastan M. M. Durga, M. Joseph Kishore, and V. Chaitanya

Abstract An unsupervised learning method based on artificial neural networks, deep learning is a branch of machine learning. Generative Adversarial Networks or simply GANs are one of the applications of deep learning methods for generative modeling. GANs are useful for generating realistic images, texts, music and many other forms of art. A lot of research work is being done in developing and improving the performance of GANs because of their potential for many real world applications. GANs are best suited for generating photorealistic images that even humans cannot identify as fake. Many of the existing models worked on only for generating image of birds and flowers from their textual descriptions. But these interesting features of GANs allow us to train the model to generate realistic images of human faces with the desired facial features. StyleGAN is one of the most popular GANs that can generate fake human faces. With the help of such models, we can try to generate images of human faces with specific facial features given as a textual description.

Keywords Generative adversarial networks · StyleGAN · Text to face generation

1 Introduction

Image generation is one of the most crucial and complex tasks in the image processing domain. Images need to be visually meaningful for the human eye. Recent years have shown drastic developments in image generation with powerful deep learning techniques. One such generative model is the GAN [1]. A GAN architecture consists of two modules, Generator and Discriminator. GANs are unsupervised learning models. In the training phase of GAN, the Discriminator acts as the supervisor for the Generator. The Generator and Discriminator are made to involve in a zero sum game. In a zero sum game, there are two players in which one is the winner and the other is loser. Incorporating this technique in the training of the GAN, the goal of the

E. Eswar Reddy (✉) · M. M. M. Durga · M. Joseph Kishore · V. Chaitanya
Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India
e-mail: eswar.eppala@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_14

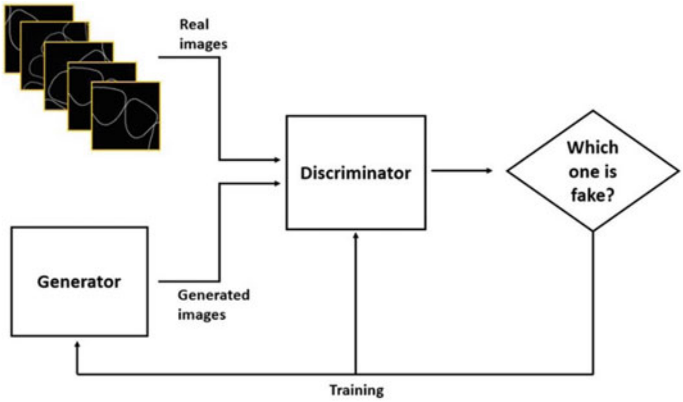


Fig. 1 GAN flow architecture

Generator is to deceive the Discriminator by generating the realistic images while the Discriminator classifies the generated images as real or fake. Figure 1 represents the architecture of the gan model.

There are state-of-the-art GAN models to generate high-resolution images with finer details. Karras et al. [2] proposed StyleGAN, such GAN architecture with a modified approach to the Generator component. StyleGAN can achieve style transfer, which is basically able to generate an image by using another image as the reference for style. With the ability to modify the properties of the generated images, StyleGAN [2] extends the capabilities of the current GAN architecture. The image is created consecutively by StyleGAN, starting at a low resolution and growing to a high resolution (1024 * 1024). In this architecture, which differs from existing GAN architectures, numerous improvements to the generator model were proposed. Without affecting other levels, it transforms the input of each level separately and analyzes the visual features that are manifested in that level, from common features (position, face shape) to minor details (hair color).

StyleGAN is a very effective approach for generating images of human faces. It was trained on FFHQ dataset [3] and has the ability to produce high-resolution images up to 1024 * 1024. But the challenge lies in controlling the facial features of the image generated in accordance with the textual description of the face. This falls under the domain of text to image synthesis. There are several existing techniques for generating images guided with textual description, but very little research work exists for text to face synthesis.

2 Related Work

Nyaupane et al. [4], proposed text to face generation involves training a text encoder, an image encoder and a DAMSM module. The text encoder generates word embeddings and sentence embedding for the given textual description. The image encoder INCEPTION V3 is used to extract the local and global features of the generated image during the generator training. In the DAMSM module, the loss of the model is calculated. Fine detailed manipulation of face features is done using region context vectors. Three stages of generators in parallel with three stages of discriminator, where the attention vector of output of each layer is fed into the next layer (Fig. 2).

Xia et al. [5] describe two techniques for text-guided variable facial picture generation and face image manipulation. A method introduced for a key component of both approaches that makes it necessary to train the image encoder before plotting real images into a GAN's latent space in order to extract all encoded information at the

Fig. 2 Text to human face images

pixel level. This gave an example of how to train a text encoder using W space's hierarchical property. Xia et al. [5] proposed visual-linguistic similarity learning, which maps the image and text into a single space. In addition, this study provides a case-level optimization that uses the learned encoder as a regularization to more precisely reconstruct the pixel values while maintaining identity throughout manipulation.

A novel method for creating text-to-face images was proposed by Xia et al. [6]. The latent code for an image is created using the StyleGAN Inversion approach. The text embeddings produced by the text encoder are then used to alter this latent code. Each of the 14 levels of the StyleGAN architecture's Generator manipulates a distinct characteristic, such as a feature's facial shape, haircut, cheekbones, or other specific attributes. This may create the required image by adjusting the latent vector at each layer in relation to the attribute in the text description. When the sampled picture is being mixed with the desired properties, the resulting image will match the text description. Instead of using the DAMSM model to map text to images, this model uses Visual Linguistic Similarity. Only the necessary picture properties can be carefully altered via instance level optimization in order to achieve results.

Two portions were proposed by Khan et al. [7]. The process of encoding text into semantic vectors will be covered in the first section, and the process of decoding text's semantic properties into accurate natural images will be covered in the second. A thorough explanation of the complete network architecture is also given. The suggested architecture's framework is built on two streams. The text is encoded in the first section, and the image is decoded in the second section using the encoded text embeddings. The textual data are transformed into a semantic vector via a text encoder. Then, using the semantic elements of the text that are encoded by the text encoder, the picture decoder produces realistic images. The majority of text to picture generating methods used by generative adversarial networks at the moment are built on the training of individual modules. The text encoder and image decoder were trained independently.

Nasir et al. [8]'s definition of their text to face problem included a description of an automatic caption generation method for the purpose of producing captions. They develop six groups of features in answer to six queries that gradually define the face, initially with the face shape to facial attributes that improve the visibility, also for the translation of the list for attributes of the human face images in the CelebA [2] dataset into respective captions. The discriminator can be expressed as D: RI RT (0, 1), Z is the size of noise vector given to the discriminator, the T is the size of the caption's skip-thought embedding, and the I is the size of the discriminator. By using skip-though encoder they encoded the textual description, after sampling the input noise, which has the dimensions Z, R, and $U(0, 1)$ (t). To reduce the size (t) to 256, of the text encodings, they have used the linked layers and leaky RELU activation. The combination of the encoding (t) and noise (Z) creates a vector of length 356. Then, this vector is given as an input to the generator. The images produced reveal encouraging outcomes. The difficulty of text to face synthesis increased due to lack of stability in training GAN and inconsistent caption length.

3 Proposed Methodology

This section proposes the our implemented methodology and architecture of our work. This section includes the modules like Text preprocessing, Random Sampling, StyleGAN inversion, CLIP loss and image generation.

3.1 Text Preprocessing

Preprocessing a text entails fundamental text cleaning procedures. These include tokenization, the elimination of punctuation, and the elimination of unnecessary words. To identify each word's meaning and comprehend the semantic meaning of the terms defining the facial characteristics, the sentence is broken up into tokens. These text descriptions need to be converted in form of word embeddings using text encoders like BERT [9] model to understand the image to text similarity.

3.2 Random Sampling

We initially sample a random image as the base for image generation. It is done by randomly sampling a latent code from the intermediate latent space of StyleGAN and giving it as the input to the StyleGAN generator to produce a face image. This image acts like an anchor point, from which the face image with desired attributes can be generated.

3.3 StyleGAN Inversion

StyleGAN inversion involves mapping a given image to the W space of StyleGAN. The random image generated in the previous step has to be mapped to W space, that is, to generate the original latent code of the random image. It can be achieved using an trained encoder that can be able to map a real image to the intermediate latent space (W) of StyleGAN. To model a simple encoder for this purpose, we use image and latent code pairs. An image generated by StyleGAN from a latent code (z^{sam}) is used to generate an image. This image is fed to the encoder to generate a latent code (z^{enc}). The difference between these two latent codes is taken as the loss and given as feedback to encoder for training. But simply training an encoder to map image and latent code is not useful. The encoder will be trained without any regard for the domain knowledge of the StyleGAN generator. To overcome this problem, an In-domain knowledge based Inversion [10] encoder is used. In this approach, the encoder is trained with real images in place of generated images. A real image

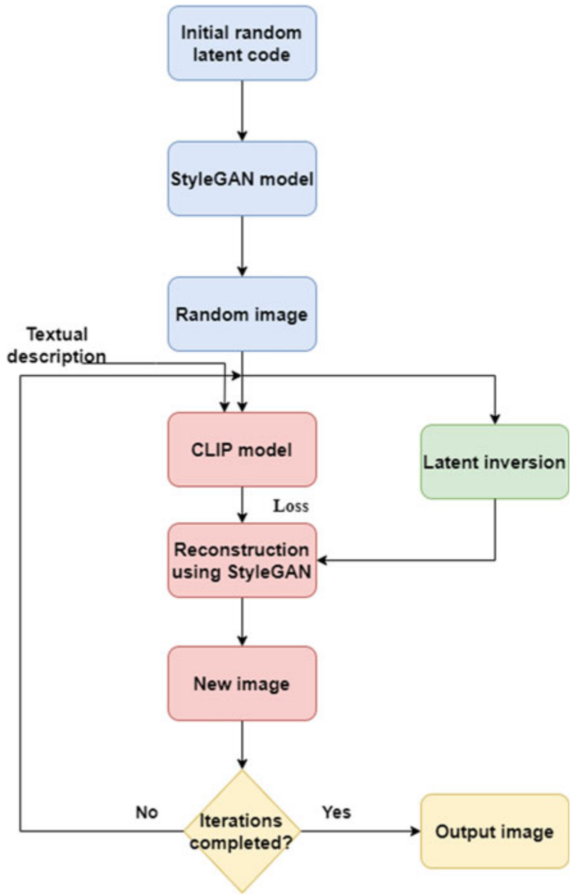


Fig. 3 Proposed system architecture

is fed to the encoder to generate latent code. This latent code is given as input for StyleGAN to generate image. The discriminator classifies the generated image as real or recreated, which is given as feedback to encoder. Figure 3 represents our proposed architecture.

3.4 CLIP Loss Function

CLIP [11] is a multimodality model that can connect an image and text. CLIP stands for Contrastive Language-Image Pre-training. This model learns the relationship between a textual description and image the text describes. It maps the image and text to a common vector space. CLIP can be useful in measuring the visual linguistic comparison between the image and the text description for the similarity. In this case, it can be used to measure how much does the generated image matches the

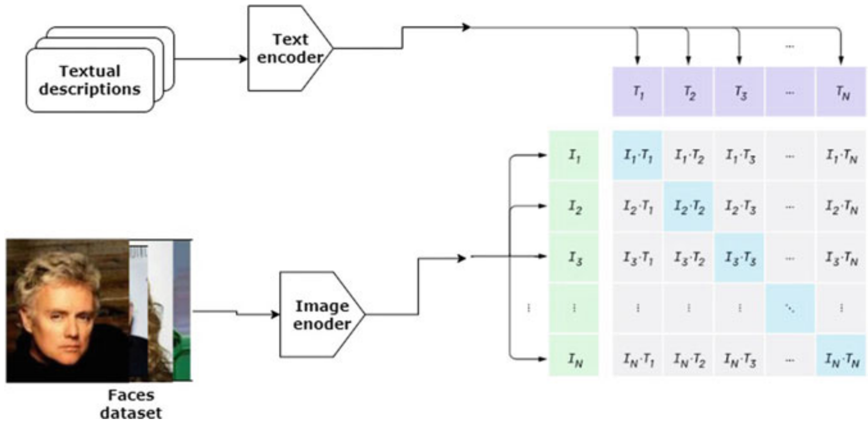


Fig. 4 CLIP model architecture

text description. Indirectly it measures how much does the image not match the text description. This value is considered in the loss function of the model along with other losses. Figure 4 shows the clip model architecture trained on the text and facial images dataset. It is a neural network architecture that can effectively pick up visual notions from natural language supervision. Similar to GPT-2 and GPT-3's "zero-shot" features, CLIP can be applied to the given input description and the generated image to measure the similarity of visual semantics and text semantics. By using CLIP model, the loss can be calculated between the generated image and the textual description. Image with minimum loss can be more accurate when campared to other images.

3.5 Image Generation

The initial random image and input text description are used to calculate the CLIP loss. This image is reconstructed by inversion and clip loss is recalculated. This process is repeated for a number of iterations to minimize the loss using Gradient Descent Algorithm.

3.6 Algorithm StyleGAN Inversion

Input: Image

Output : Latent code

Step-1: Sample a real image (xreal) from the data distribution

Step-2: Input the image (xreal) to the encoder to generate latent code ($z1$)
Step-3: Generate another image (xrec) using StyleGAN with this generated latent code (z1)
Step-4: Calculate the loss function using these two images xreal and xrec
Step-5: Repeat above steps while minimizing the loss to train the encoder

Image Generation

Input: Textual description

Output: Face Image

Step-1: Sample a latent code and input to the Stylegan to generate random image
Step-2: Initialize adam optimizer
Step-3: Input the text description and generated image to CLIP to calculate the loss
Step-4: StylegGAN inversion module is applied on the image to generate latent code
Step-5: Image is reconstructed by giving latent code as input to StyleGAN
Step-6: Repeat step 3 to 5 for fixed number of iterations.
Step-7: Output the final generated image.

4 Result and Analysis

Below are the some of the results we have obtained. The textual description is given as the input to the model. The Generated images are produced by reconstruction and loss minimization through a number of iterations. The generated images show promising results of this approach. The accuracy and quality of the image increases with increasing the number of iterations (Fig. 5).

There are two metrics for evaluating the performance of images generated by GAN. Frechet Inception Distance is most popularly used evaluation indicator for gauging the effectiveness of GAN-generated images. The higher the image quality, the lower the FID score Perceptual Image Patch Similarity Training (LPIPS).

The perceptual similarity between two images is assessed using LPIPS. This statistic enables us to comprehend the variety of the images the model produces. The various models' accuracy is examined.

FID (Frechet Inception Distance) is a standard evaluation metric quality measurement of GAN-generated images. A Low FID Score indicates the model generates better quality images. LPIPS (Learned Perceptual Image Patch Similarity) is used to judge the visual similarity between two generated images. This metric allows us to understand the diversity of the images generated by the model.

Based on Adam optimizer, clip loss can be calculated, it can be used to evaluate the accuracy and loss (Figs. 6, 7 and 8; Table 1).

Text Description	Generated Image
The woman has black hair, big nose, mouth slightly open, wavy hair, and bushy eyebrows. She is young, and smiling and wears lipstick.	
This young, and attractive man has bags under eyes, bushy eyebrows, big nose, black hair, and high cheekbones.	
This person has big nose, mouth slightly open, high cheekbones, rosy cheeks, arched eyebrows, and oval face.She is smiling, and young.	

Fig. 5 Output results

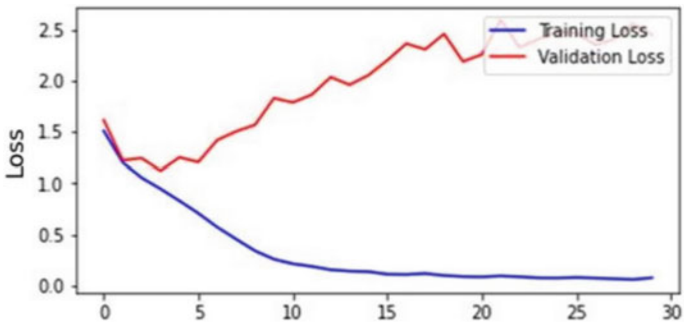


Fig. 6 Model loss for Adam optimizer

5 Conclusion and Future Work

In this approach we try to make use of two different models namely CLIP and StyleGAN. The main objective is to find the optimal method that can generate the human face images. The generated images show the capability of manipulating the intermediate latent space of StyleGAN to generate desired images, but with not very high accuracy. This is due to the initial random sampling of latent code. The image produced at the end of this generation pipeline highly depends on the initial image.

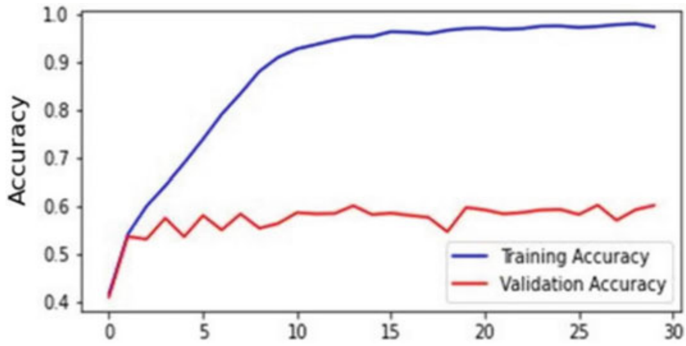


Fig. 7 Accuracy graph for Adam optimizer

Fig. 8 Adam optimizer

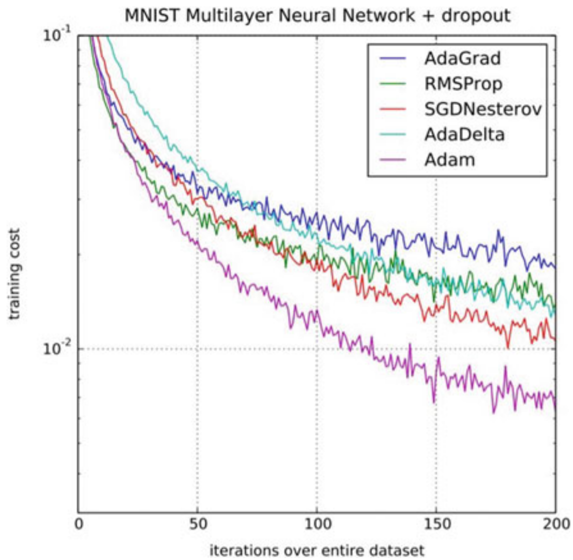


Table 1 Accuracy and time taken table for the different number models

Model	FID score	LPIPS	Accuracy
AttGAN	120.98	0.512	14.2
ControlGAN	116.32	0.522	18.2
Our model	106.37	0.456	25.3

There is also the problem of feature entanglement, where some facial features are naturally entangled and accurately manipulating specific features without changing others is crucial. There is scope for future work in overcoming these drawbacks.

References

1. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2020) Generative adversarial networks. Commun ACM 63:139–144. https://doi.org/10.1145/3422622
2. Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)
3. Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. In: CVPR, pp 4401–4410
4. Nyaupane BK, Tamrakar J, Nyaupane BK (2021) Synthesizing human face image from textual description of facial attributes using attentional generative adversarial network. In: 10th IOE graduate conference, vol 10, Oct 2021
5. Xia W, Yang Y, Xue J-H, Wu B (2021) Towards open-world text-guided face image generation and manipulation. In: IEEE conference on computer vision and pattern recognition, 18 April 2021
6. Xia W, Yang Y, Xue J-H, Wu B (2021) TediGAN: text-guided diverse face image generation and manipulation. In: CVF conference on computer vision and pattern recognition, CVF conference on computer vision and pattern recognition (CVPR)
7. Khan MZ et al (2020) Realistic image generation of face from text description using the fully trained generative adversarial networks. IEEE Access 9
8. Nasir OR, Jha SK, Grover MS, Yu Y, Kumar A, Shah RR (2019) Text2facegan: face generation from finegrained textual descriptions. In: IEEE fifth international conference on multimedia big data (BigMM), 26 Nov 2019
9. Devlin J, Chang MW, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding
10. Zhu J, Shen Y, Zhao D, Zhou B (2020) In-domain GAN inversion for real image editing. In: ECCV
11. Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, Krueger G, Sutskever I (2021) In: Proceedings of the 38th international conference on machine learning, PMLR 139:8748–8763

Rumour Detection Model for Political Tweets Using ANN

Mohammed Mahmood Ali, Mohammad S. Qaseem, and Syed Shabbeer Ahmad

Abstract Evolution of technology, especially in online messaging applications, has evolved the communication era. Trending informative social messages has become a part of lives of people with its dynamic transferring capabilities. The prominent organizations profiles are identified by people based on their organizational profiles uploaded in social media. Political goons started to smartly make use of various social media platforms (e.g., Twitter, Facebook, LinkedIn, Snapchat and many more) to propagate their beneficial and welfare ideologies to attract innocent voter's attention to win the elections. False promises and political rumours made by these selfish politicians have been a serious threat for this society. Many of researches depend on topic-independent political rumour detection on Twitter which has been well-known by everyone. In this paper, we proposed a rumour detection algorithm that has the ability to monitor political rumour messages based on tweets. In this paper, we have deeply surveyed recent methods developed to identify political rumours from tweets and found around eight (8) papers exclusively. Twitter usage had drastically grown up for the use of political rumours apart from WhatsApp and Facebook social sites. It is found that usage of artificial neural network for identifying political tweets is better when compared to rainforest, decision trees and KNN.

Keywords Twitter social network · Rumour detection system · Sample stream and filter stream · Artificial neural network (ANN)

M. M. Ali (✉) · M. S. Qaseem
Muffakham Jah College of Engineering and Technology, Hyderabad, India
e-mail: mahmoodedu@gmail.com

S. S. Ahmad
Nawab Shah Alam Khan College of Engineering and Technology, Hyderabad, Telangana, India
e-mail: shabbeer.ahmad@mjcollege.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_15

1 Introduction

1.1 Background Knowledge

Few of the selfish political organizations started to use the various social media platforms (e.g., Twitter, Facebook, LinkedIn, Snapchat and many more) to propagate their false development and welfare ideas to get the voter's interest to win elections. The advanced messaging system that gets acknowledged by social platforms from the population had boosted a biggest change in the history of communication era. The tactics and notifications of messaging system in shortcuts led the people's life easier and safer by saving lot of time in conveying of useful information and news as well. Subsequently, fake and deceitful information also is on edge of the gash in social media, especially political news which is indirectly a worrying etiquette for the society as it hinders the economic development plans, which needs stringent action by the government in the secular aspects [1]. The objective is to trap untruthful political rumours from tweets which are transmitted through innumerable social messaging apps. Apart from catching those online social criminals who intentionally or unknowingly transfer such fake political rumours are to be identified and informed to the security surveillance department for legal action [2].

1.2 Problem Statement

Fake news of political rumours especially in social networking arena gives a very bad impact of influencing the innocent people and distracting the mental peace as well as harmony of the nation. There should be spying centralized controlling system that controls those political rumours by keeping a stringent monitoring eye on such microblogs which are sent via the social networking systems. Instant messages before reaching the online social users community, especially pertaining to political aspects, must be automatically searched in authorized political databases which needs to be implemented by enhancing the security. Unless, until the political news is authentic, then only these servers permit those political news to other online social users; otherwise straightforwardly the servers deny the tweets. Further, the servers automatically need to generate an alarming system to the spying department informing the details of online tweeter who tried to initiate annoyance in the nation for political benefits. Among the false tweets of political rumours, monitoring is not an easy task to catch hold for them, which requires extreme processing tasks to block which is not an easy task to check its authenticity [3].

Political rumour detection system is required to handle not only one specific topic of political rumour category, but it needs to identify and pinpoint the different classes of political tweets/microblogs from Twitter. The following problem is analysed based on the requirement and depicted below:

- **U** set of user on a social media network $u_i \in U$, u_i is each user in Twitter.
- **Political rumour state (pr)**: (pr, u_i, t) or $\mathbf{pr}^t$ denotes rumour state of user u_i.

$$pr_i^t \in \{0, 1\} \text{ if } r_i^t = 1 \text{ is rumour detected and}$$
$$\text{if } pr_i^t = 0 \text{ is rumour not detected}$$

- **Political rumour state**: It is series of **T** partially labelled graph in Twitter

$$\left\{G^t = \left(U^t, M^t, R^t\right) | (t = \{1, 2, \ldots T\})\right\}$$
$$f = \{G_u^1, G_u^2, \ldots G_u^t\} \rightarrow \{R^1, \ldots R^T\}$$

the objective is to learn a function to predict unlabelled users rumour states.

The idea is to build a rumour detection model, which is depicted in the above function(f), from social media microblogs tweets sent by the users in the social network such that when rumour is found it has to be predicted and notified [2].

2 Literature Survey

Free messaging application systems would infinitely influence the Internet communications and badly affect the social life. The political parties started to utilize this communication era to gain political goals and shape their cunning careers with the use of fake political rumours. They willingly propagate misleading information in the way of news to win people hearts in multiple ways. Further, there are a slew of websites devoted only to broadcast deceptive information, which is untruthful, through publicity materials, manufacturer products and fear-mongering ideas under the pretext of real news. The main goal of fake news via websites is to influence public opinion on specific topics (*to build political career*). This may be mostly seen in foreign countries such as Ukraine, USA, UK, Russia and also in multiple countries. As a result, fake news is a worldwide issue and a vast challenge to deal with especially in politics. The survey carried out in political rumour domain is presented in Table 1 that gives the summary of how political rumours presently exist and the technique devised to resolve is elaborated in Sect. 3. The datasets can be collected from PEER Journal, pertaining to various political rumours and tried our best to tackle such rumours and proposed the political rumour detection strategy especially for Twitter [2, 4–6].

Table 1 Political rumour models developed earlier and their drawbacks

S.No.	Title	Objective	Strategy/ technique	Remarks
1	Analysing the methods of rumour detection in social media [4]	Misinformation sent is named as rumour in the social media. The focus of this paper is to predict the rumours from collectively embedding machine and deep learning methods through training them by datasets	Various machine learning, deep learning and hybrid models are used collectively in one system to identify rumours	It is time taking process as it requires training with huge datasets
2	Covid-19 health checking related rumour and misinformation in Twitter [1]	Build using partial features by embedding potential as well as network related to Covid-19 health problems from specific domain for various types of diseases	Pipeline method of the rumour detection process	The accuracy of 89% when tested on real datasets is identified
3	Cross-topic of rumour identifying in health domain [7]	Rumour identifying from messages at different levels by gathering tweets, in which very less effort is used to check the rumour by picking individual tweet	Strength of the parameter set is taken from earlier works as a pipeline	Crossing involves only first part as a test set and other one as a training in the experiments, which is not so precise
4	Rumour detection using programming from microblogs [8]	Preprocess the existing micro blogs forwarded via various social network	Machine learning technique is employed to predict such tweets	Unable to identify long and multimodal spam messages forwarded in social sites
5	Big Data and eminence for false and improper information [9]	The issues involved are gathering of qualitative data of untruthful and genuine fake news based on equalized distribution of topics	NLP uses the parametric approach, which constitutes of extracting and analysing of linguistic for identifying of a particular target phenomenal messages	To trace fake information, it needs bulk amounts of datasets for training

(continued)

Table 1 (continued)

S.No.	Title	Objective	Strategy/ technique	Remarks
6	Aspect stage sentiments analysing using NLP and deep methods [5]	Sentiment is a NLP technique that can be found in comment or tweet. By using sentiments of various users exclusively based on a particular aspect, extraction is done by employing of deep learning methods	CNN is used for image classification, and an NLP is used for processing of text, i.e., sentiment analysis	CNN method also sometime fails to identify valid aspects
7	Dynamic features-based rumour detection method [10]	In these extraction of word, symbol and emotion features of rumour remark as dynamic feature due to the remark information which expresses the change in rumour spreading	Feature change extraction framework (FCEF) for handling rumour incident with the help of two detection method neural network and machine learning	The accuracy of 90% when tested on real datasets
8	Applying of supervised AI algorithms for fake news [11]	Parameter extraction is used on false rumour news dataset by removing few columns from the existing parameters. For this, assisted AI algorithms are employed	Employed text dredging and assisted AI algorithms in social media on fake news	Improvement to these can be achieved by discovering novel algorithms by joining existing algorithms for accurate precision rates

3 Proposed Rumour Detection System Using ANN

The rumour detection model is proposed using artificial neural network (ANN). The steps involved in the algorithm-cum-flowchart are shown in Fig. 1, which are elaborated as follows:

Step 1: Input the data extracted from Twitter API for training purpose.

Step 2: In this sequence, preprocessing steps are involved for extracting tokens (words) by removing stop words, hash tags and applying stemming algorithm.

Step 3: Accumulate all the words and cross-check for spelling correction and validate.

Step 4: Train using ANN and build the model:

(i) Assign a random weight to the start of the algorithm.

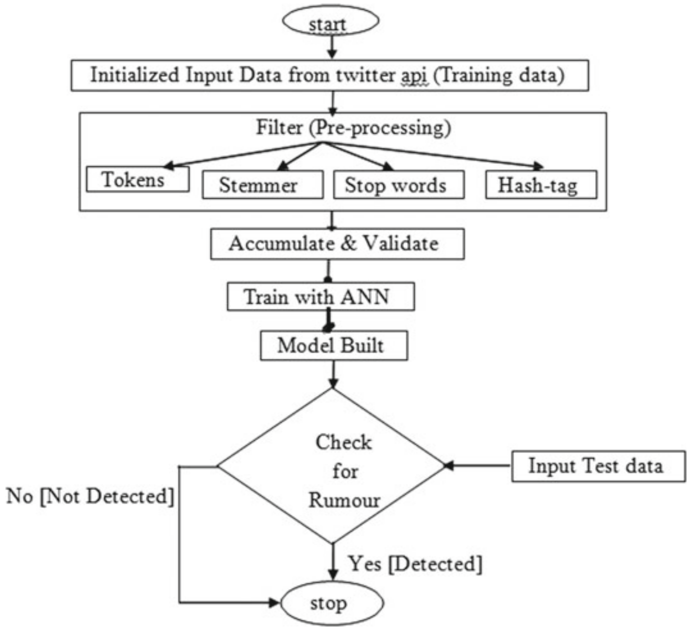


Fig. 1 Suggested rumour detection system by embedding artificial neural network

(ii) Determine the activation rate of the hidden node using the input and linkage.
(iii) Finding of the activation rate of output node using the node and links.
(iv) Calculate the error rate.
(v) Cascade the error down to the hidden node using the weights and error discovered at the output node.
(vi) Weights between hidden and input node should be recalculated.
(vii) Repeat the procedure until the criterion is fulfilled. Score the activation rate of the output node using the final weights.

Step 5: Check of rumour or non-rumour by giving test data
Step 6: If "detected" report it and stop, or else stop saying "not detected".

4 Conclusion and Future Scope

Towards social networks, there has been a lot of researches in the field of rumour detection. However, mediocre studies have seriously worked on tweets of Twitter. Analysing datasets of Facebook consists of many restrictions, such as data access permission owing to user data privacy and, as a result, a lack of high-quality data. By demonstrating the spread of social media in all spheres of life, we highlight the necessity for new approaches to rumour identification that target rumours other than political rumours [12]. Exclusively, the third largest smaller-scale blogging platform, i.e., Twitter, is found to be grown drastically for the use of political rumours apart from WhatsApp and Facebook. In this paper, we critically surveyed eight (8) papers presented in Table 1 that explored how fake rumours are identified from various social networking apps and found that artificial neural network (ANN) will be the best among all of other techniques. At this specific point, we proposed ANN rumour detection model in Sect. 3 that will efficiently start investigating the tweets to predict fake political rumours.

Future work is to consider ideology of the political goons with respect to their behavioural and psychological aspects generated by them through microblogs shared via WhatsApp, LinkedIn and Facebook apart from the tweets. Successively, perform the rigorous checking of replies of those users in Twitter who responded to such rumour tweets which may hinder the peace of society, from all over the world bringing them into one unique surveillance platform of social media [13, 14].

References

1. Gautam KS, Dirkson A, Majchrzak TA (2021) An exploratory study of COVID-19 misinformation on Twitter. Online Soc Netw Media J 22:201–218
2. Mahmood MA, Lakshmi R (2013) Framework for surveillance of instant messages. Int J Internet Technol Secur Trans 5(1):18–41. Inderscience
3. Ali MM et al (2019) Rumour detection models & tools for social networking sites. Int J Eng Adv Technol (IJEAT) 9(2):3291–3296
4. Ajeet RP, Mahajan A, Keshav S, Aishwary P, Anusha N (2020) Analysis of techniques for rumor detection in social media. In: Procedia computer science, vol 167. Elsevier, pp 2286–2296
5. News Use Across Social Media Platforms in 2020. Pew Research Center (2021) https://www.pewresearch.org/journalism/2021/01/12/news-use-across-social-media-platforms-in-2020/. Online Accessed on 21 Nov 2021
6. Arkaitz Z (2021) Fake news detection: a survey of evaluation datasets. Peer J Comput Sci J. National Library of Medicine
7. Ali G, Malik MSI (2022) Rumour identification on Twitter as a function of novel textual and language-context features. Multimedia tools application. Springer
8. Zongmin L, Qi Z, Xinyu D, Yanfang M, Shihang W (2021) Social media rumor refutation effectiveness: evaluation, modelling and enhancement. Inf Process Manag 58(1)
9. Thi Dieu Linh N, Lu Z (2021) Eminent role of machine learning in the healthcare data management. In: Data science and medical informatics in healthcare technologies. Springer, briefs in applied sciences and technology. Springer, Singapore
10. Meng Z, Yu SR, Li Jiang G, Song Y (2020) Dynamic features based rumor detection method. In: 2020 Chinese control and decision conference (CCDC). IEEE, pp 379–384
11. Pandey S, Gupta R, Kumar J (2022) Identification of fake news using machine learning techniques. In: Pundir AKS, Yadav N (eds) Recent trends in communication and intelligent systems. Algorithms for intelligent systems. Springer
12. Mahmood MA, Qaseem MS, Rasheed MA, Khan MKA (2020) ESMD: enhanced suspicious message detection framework in instant messaging applications. In: Fourth international conference on inventive systems and control (ICISC), pp 777–784

13. Ali MM, Hajera S (2018) Psychological stress detection from social media data using a novel hybrid model. Int J Intell Syst Appl Eng 6(4):299–305
14. Ali MM, Noorain S, Qaseem MS, ur Rahman A (2022) Suspicious human behaviour detection focusing on campus sites. In: Chaurasia MA, Juang CF (eds) Emerging IT/ICT and AI technologies affecting society. Lecture notes in networks and systems 478. Springer

Face Detections of a Person Wearing COVID Mask Using Multi-task Cascaded Convolutional Neural Networks

Ramana Naagavelli, Puppala Tirupathi, and Akhil Kumar Yadanala

Abstract According to a report, over one percent of the world population is infected with COVID-19 in the last year. To avoid the spread of the infection, World Health Organization (WHO) suggested individuals to wear a face mask and maintain social distance. The use of face masks raised concerns about the accuracy of the facial recognition system used for office attendance, unlocking phones, etc. Masked faces make it difficult to recognize the individual. The objective of this study is to figure out who is wearing a face mask. There are two steps in recognizing a person's identity; they are face detection and face recognition. The faces are detected using multi-task cascaded convolutional neural networks (MTCNN) trained on the WIDER FACE dataset. Face descriptors were obtained from MTCNN's suggested region of interest using the model VGGFACE2. The MS-Celeb-1M dataset was used to train the VGGFACE2. Using facial descriptors, it is possible to identify a person. A cosine distance criterion is utilized to verify individuals, and the K-nearest neighboring classifier is used to determine a person's identity. This research will provide better results in identifying a person wearing a face mask.

Keywords CNN · COVID-19 · Face recognition · KNN

1 Introduction

1.1 Face Detection

Face detection is a term that refers to computer technology that can detect the presence of people's faces in digital images [1]. The face detection model is later used in the application like face recognition.

R. Naagavelli (✉) · P. Tirupathi · A. K. Yadanala
Department of CSE, Kakatiya University, Hanamkonda, India
e-mail: ramanauce.ku@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_16

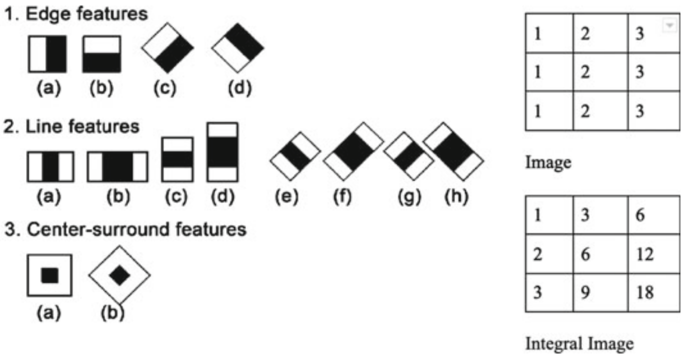


Fig. 1 Haar-like features

Face detection techniques

There are various methods for detecting faces, but the Viola–Jones technique [2] is the most often used. The following is a comprehensive summary of Viola–Jones. One of the most commonly used approaches for detecting faces is Viola–Jones. Viola–Jones employs Haar-like features, integral images, AdaBoost, and the cascade classifier to produce quick and accurate results.

Haar-like features

Viola–Jones has the Haar-like features and among them the most popular features are line features, edge features, four-sided features, and a few others which are the most often utilized. These features have black and white regions. These Haar-like features slide over the image to find the facial features. The difference between the total of black pixels and white pixels is used to determine the value of the characteristics of the face. Figure 1 shows Haar-like features.

Integral images

Integral images is a software program that allows you to rapidly do complex computations. The total of all pixels above and to the left, including the target pixel, is calculated in integral pictures to get each point.

AdaBoost

AdaBoost was the first boosting method used in binary classification. AdaBoost is a machine learning technique that uses a large number of weak classifiers to produce strong classifiers. In a 24×24 candidate window, there are 160,000 features, which are given as input to AdaBoost, which then determines the strong classifier and the weak classifier.

Cascading classifiers

It returns the finest characteristics after executing AdaBoost. Then, we put those characteristics into a cascade classifier, which quickly removes features that are not

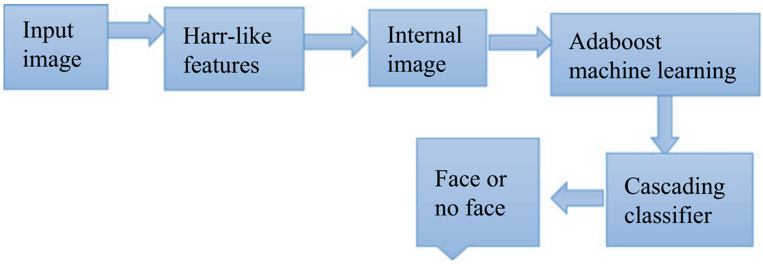


Fig. 2 Work flow of Viola–Jones

linked to the face. We divided the cascade process into many phases, with each stage attempting to locate the face's characteristics (Fig. 2).

2 Convolution Neural Network

A convolutional neural network (CNN) [3] is a deep learning approach for processing images. CNN has a three-step procedure: (1) convolution, (2) pooling, and (3) fully connected layer. The initial layer of a neural network is convolutional. Convolutional is a method of utilizing filters to preserve the relationship between pixels. One of the matrices considered as an image matrix, while the other is the kernel matrix. To construct a feature matrix, perform mathematical operations on both the image and kernel matrices. The most commonly used filter is edge detection, horizontal line detection, image blurring, image sharpening, etc. This procedure is carried out with a stride 1. Stride is simply a shift in the number of pixels across the input images (Fig. 3).

Pooling is the technique of down sampling a feature map from higher dimensions to lower dimensions without losing the image's feature. There are three types of pooling: maximum pooling, average pooling, and some pooling. In max pooling, it selects the large number within the kernel.

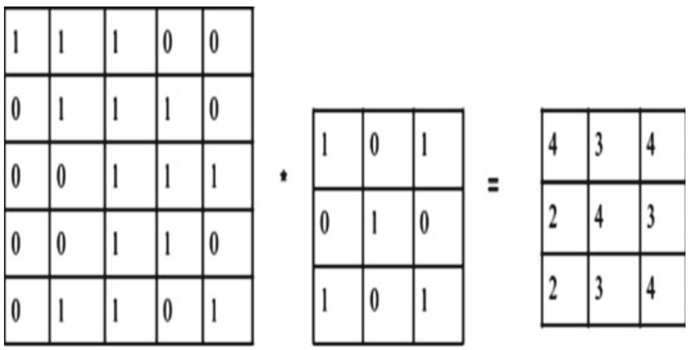


Fig. 3 Convolution

3 Proposed System

People used face masks during the COVID epidemic, which created significant issues in face recognition. Is it feasible to build a model that uses both masked and unmasked faces as input and predicts the identification of unknowns? Faces may be detected in a variety of ways. The Viola–Jones face detection model is well known and frequently utilized. Face recognition may be achieved with several pre-trained models and approaches including face matching, face similarity, and face transformation.

The techniques described above might lead to the problems outlined below. Viola–Jones creates a sliding window and searches the candidate window for a face. It is looking for characteristics that are Haar-like features. But Viola–Jones can detect the front part of a person's face but not when the person poses in the upward, downward, or sideways direction, and in few cases, it becomes difficult to detect the person's face when the face is covered with a face mask. Face recognition can be achieved using the various pre-trained model and other approaches, but most of the pre-trained models are trained on low-resolution images so when a person wears a facemask it becomes difficult to extract the facial features which lead to wrongly predicting the individual and also there is no reliable dataset for face recognition of masked faces.

4 Experimental Results

To forecast the individual, one way is to use threshold values on cosine distance, while the other is to use the KNN machine learning algorithm. After numerous tests, it was discovered that a cosine distance of 0.7–0.8 works well, and it was also suggested in the FaceNet [4] study. However, for every input image, a lot of computation is to be done on the entire database, and it takes a lot of work and time. KNN may be used to speed things up while also improving accuracy. So, with accuracy in mind, we are considering the KNN machine learning method for the prediction of individuals. Figure 4 shows the output images while performing the experiment (Fig. 5).

Using KNN, a person is predicted based on the data points surrounding it. Is it feasible to anticipate the K values for a given dataset? This may be accomplished by carrying out several experiments. With the dataset, we created many sets. The first set contains 100 subjects, each with 5 images, the second set contains 100 subjects, each with 10 images, the third set contains 100 subjects, each with 15 images, the fourth set contains 100 subjects, each with 20 images, the fifth set contains 100 subjects, each with 25 images, and the sixth set contains 100 subjects, each with 30 images. The minimum and maximum number of images per individual are 20 and 30, respectively. In all trials, assume that the test dataset set contains 56 masked images (Tables 1 and 2).

When the model was trained on 100 different subjects, each with 5 images at K equal to 9, the greatest accuracy score attained is 0.82. When the model was trained

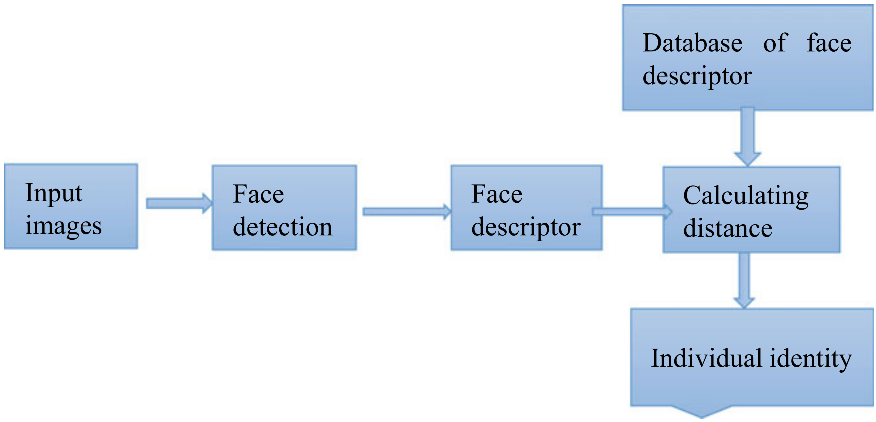


Fig. 4 Proposed system workflow

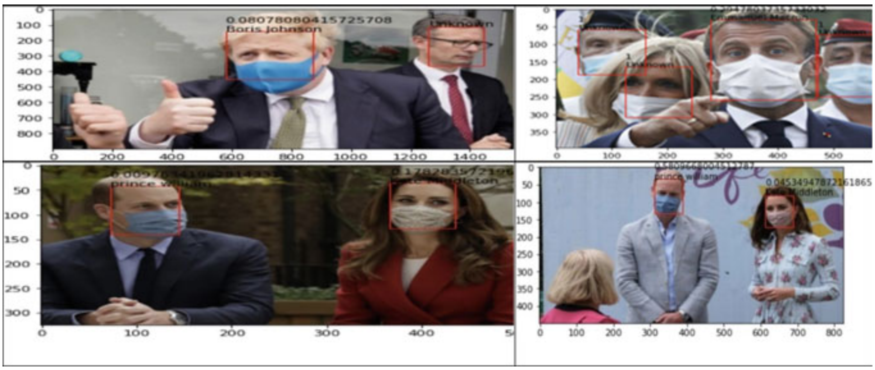


Fig. 5 Masked image input

Table 1 Masked images as input

Trained copies per subject	*K*	Accuracy
5	9	0.82
10	9-13-17-19	0.78
15	15-17-19-21-23-25-27-29-31	0.76
20	21-23-25-27-29	0.78
25	21-23-25-27-29-31-33-35	0.78
30	27-29	0.80

Table 2 Unmasked images as input

Trained copies per subject	*K*	Accuracy score
5	1-3-5-7-9-11	0.96
10	7-9-11-13-15-17-19-21-23-25-27-29	0.98
15	3-5-7-9-11-13-15-17-19-21-25-27-29-31-35	0.94
20	1-3-5-7-9-11-13-15-17-19-21-25-27-29-31-35-37	0.95
25	1-3-5-7-9-11-13-15-17-19-21-25-27-29-31-35-37-39-41-43-45-47	0.95
30	1-3-5-7-9-11-13-15-17-19-21-25-27-29-31-35-37-39-41-45-47-49-51-53-55-57-59-61	0.95

on 100 different subjects, each with 10 images at K equal to 9-13-17-19, the greatest accuracy score attained is 0.78. When the model was trained on 100 different subjects, each with 15 images at K equal to 15-17-19-21-23-25-27-29-31, the greatest accuracy score attained is 0.76. When the model was trained on 100 different subjects, each with 20 images at K equal to 21-23-25-27-29, the greatest accuracy score attained is 0.78. When the model was trained on 100 different subjects, each with 25 images at K equal to 21-23-25-27-29-31-33-35, the greatest accuracy score attained is 0.78. When the model was trained on 100 different subjects, each with 30 images at K equal to 27-29, the greatest accuracy score attained is 0.80.

4.1 Observation

The average accuracy score for masked faces is 0.78. Unmasked faces, on the other hand, have an average accuracy score of 0.99. As per the experiment, K value should be the average number of images per subject. The other approach of face recognition was tested using seven invariant moments, and the accuracy score was below 0.4. I tested my hypothesis using the LFW synthetic dataset. It achieved the maximum accuracy of 0.44 at $K = 3$. We began investigating why there was a drop in inaccuracy when we saw it. We gathered the results of the above experiment's mistake. The input data point is surrounded by the data points of other subjects when we applied the KNN method. The observation made from above is that there is a possibility of a masked face matching with another subject.

5 Conclusion

The primary focus of this project is face detection and recognition of a person wearing a face mask. Due to a rise in COVID-19, it has become difficult to recognize the individual. In recent years, improvements in image processing technology became possible to tackle the problem in the real world. As of now, there is no reliable dataset for the recognition of masked faces. So, the only approach is to create a synthetic dataset. The synthetic dataset is created by using the Dlib machine learning tool which plots the coordinates on the face, and these coordinates help in overlay the face mask. Due to the increase in technology, it is possible to create a real-world dataset from the Internet using Google programmable search engine and Google image API. In future, to create a large-scale dataset for masked face recognition with higher resolution, to create a lighter version of this project, so it can run on lighter devices such as mobile and Raspberry Pi.

References

1. Nabatchian A, Abdel-Raheem E, Ahmadi M (2008) Human face recognition using different moment invariants: a comparative study. In: 2008 congress on image and signal processing
2. Huang GB, Ramesh M, Berg T, Learned-Miller E (2007) Labeled faces in the wild: a database for studying face recognition in unconstrained environments. Technical report
3. Zhang K, Zhang Z, Li Z, Yu Q (2016) IEEE. Joint face detection and alignment using multi-task cascaded convolutional networks. Cite as: arXiv:1604.02878 [cs.CV]
4. Schroff F, Kalenichenko D, Philbin J (2015) FaceNet: a unified em-bedding for face recognition and clustering. Cite as: arXiv:1503.03832 [cs.CV]
5. Young M (1989) The technical writer's handbook. University Science, Mill Valley, CA
6. Parmar DN, Mehta BB, Face recognition methods applications. Cite as: arXiv:1403.0485 [cs.CV]

Classification of Fruit Essential Oils Using Machine Learning Practices

Katta Subba Rao, Sirisha Potluri, S. Venkateswarlu, and Madhu Bandari

Abstract Fruit essential oils perceive a lot of demands in various industrial meadows such as medicines, cosmetics, and foodstuff. Adulteration of essential oils results in replacing its high price ingredients with low-priced and quality ingredient replacements. These products can be eye-catching and profitable but observe complications and side effects on consumers. Hence, it is essential to apply an efficient novel practice to authenticate the quality of fruit essential oils and their percentage of adulteration. The nonvolatile portion of fruit essential oils is usually undervalued due to their trivial influence on the smell profile. The purpose of our research is to outline various machine learning practices that let authenticate fruit essential oils by chromatographic procedures. Naïve Bayes (NB), Neural Networks (NN), Multi-layer Perception (MLP), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Logistic Regression (LR) techniques are compared, and their performance metrics are studied to propose best suitable practice for essential oil classification.

K. S. Rao
Department of Computer Science and Engineering, B V Raju Institute of Technology, Narsapur, Medak (District), Telangana 502313, India
e-mail: subbarao.k@bvrit.ac.in

S. Potluri (✉)
Department of Computer Science and Engineering, Faculty of Science and Technology (IcfaiTech), ICFAI Foundation for Higher Education, Hyderabad 501203, India
e-mail: sirisha.potluri@ifheindia.org

S. Venkateswarlu
School of Technology, Woxsen University, Kamkole, Sadasivpet, Hyderabad, Telangana 502345, India
e-mail: registrar@woxsen.edu.in

M. Bandari
Department of Data Science and Artificial Intelligence, Faculty of Science and Technology (IcfaiTech), ICFAI Foundation for Higher Education, Hyderabad 501203, India
e-mail: madhu.bandari@ifheindia.org

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_17

Keywords Fruit essential oil · Classification · Prediction · Machine learning · Authentication

1 Introduction

Fruit essential oil has a great demand in various industrial domains due to its significance in the preparation of juices, slices, extracts, and other essential products. Fruits are vastly cultivated everywhere in the world with a wide range of verities and statistics revealed that bananas, watermelons, apples, oranges, grapes, mangoes, and guavas have significant yield and production. Fruits are classified into drupes, domes, and citrus [1]. Drupe fruits contain a single seed using which a new plant will grow. Examples of drupes are peach, mango, plum, cherry, etc. Dome fruits contain many seeds using which new plants will grow. Examples of domes are apples, papaya, pears, quince, etc. Citrus fruits contain leathery skin or layers with many juicy segments or pellets. Examples of citrus are lemon, orange, grape, lime, etc.

Essential oils are considered to be volatile and fluid aroma amalgams obtained from natural bases, typically plants [2]. Even though they are not considered to be oils in a strict nous, but exhibit poor solubility in water. These natural essential oils usually have an aroma and are thus applied in food seasoning and perfumery. Through various fragrance extraction procedures such as distillation, solvent extraction, and cold pressing, we extract these essential oils. Naturally, essential oils are extremely complex combination compounds and are perceived as hundreds of specific aroma mixtures. Essential oils are obtained from various citrus fruits, namely orange, lemon, and grapefruit [3]. Different parts of citrus fruits are used as bases for essential oil isolation such as mesocarp, endocarp, carpel, and exocarp. To extract and distribute the main components of these natural oils, we use solvent-free microwave extraction and diffuse reflectance infrared Fourier transform. Major components of citrus essential oils contain limonene (major constituent with 50–80%), β-myrcene, and γ-terpinene [4, 5]. Lemon oils provide high quantities of oxygenated mixes, whereas orange and grapefruit are richer in citral isomers. Commonly, fruit essential oils extracted from the exocarp layer that contains rich oxygenated monoterpenes; however, those oils that are extracted from the exocarp and mesocarp layers are richer in flavone and furanocoumarin derivatives [6, 7].

2 Essential Oil Extraction Methods

Essential oils from fruits are used in an extensive range of consumer things—detergent cakes, soaps, cosmetic products, medicines, perfumes, confectionery items, soft/hard drinks, and insect repellents. The world's manufacture and consumption of fruit essential oils are growing very fast. Production technology is a vital component to increase the global yield and quality of fruit essential oil. These oils are acquired from raw materials of plants by using several extraction techniques [8–10].

2.1 Classical and Conventional Essential Oil Extraction Methods

There are several classical and conventional essential extraction methods that are still in practice throughout the world. Hydrodistillation, steam distillation, solvent extraction, Soxhlet extraction, and cold pressing method are generally used methods to extract essential oils from parts of a plant [11–13].

2.2 Non-traditional Essential Oil Extraction Methods

Classical methods of essential oil extraction have been getting replaced with innovative and non-conventional techniques due to technological advancement. Supercritical fluid extraction, microwave-assisted hydrodistillation, ultrasound-assisted extraction, solvent-free microwave extraction, and microwave hydrodiffusion and gravity are generally used as innovative and non-conventional methods to extract essential oils from parts of a plant [14–17]. The essential oils are found in fruits of various plants and are extracted using a treatment with comminution, temperature, water, and chemical solvents to obtain the outcome. The significant fundamental processes used for extraction of essential oil are concentration, solvent extraction, and automatic expression, with developments or modifications hosted for each when available. Resinoids (extracted from resin), concretes (extracted from flowers and other plant parts), absolutes (extracted from aromatic plants), and other byproducts are expressively used all over the world.

3 Quality Assessment of Essential Oils

3.1 Factors and Parameters

Several factors, overall market situations, geographical obtaining, and unexpected weather happenings, can change the quality of the essential oils each year. These are analogous apprehensions for various natural yields when supply is reserved. There is a number of parameters that we consider to ensure the quality of the essential oils.

3.2 Gas Chromatography and Mass Spectrometry

It is a practice that separates the constituents of an essential oil rendering to the variances in molecular capacity. It delivers a pattern of the oil and discloses additives, but it is not constantly convincing. Gas chromatography identifies the existence and relative measures of chemical elements in essential oil. A mass spectrometry detector let us to identify of chemical composition profile of essential oils. The gas chromatography and mass spectrometry components' summary does designate whether the essential oil is from a genuine source, of substandard quality, or if it is tainted with other, less affluent essential oils, cologne mixtures, plasticizers, or other annoying components.

4 Machine Learning Practices for Essential Oil Classification

To perceive the incidence of chemical composition profile in essential oils during its processing stage, machine learning models are used. By assessing the retention index of gas chromatography, a machine learning model can make the predictions. Various machine learning techniques, Nearest Neighbor, Multi-layer Perception, Decision Tree, Naïve Bayes, Regression, Neural Networks, and Support Vector Machine, are used to classify essential oils [18–22]. EssOilDB database with 123,041 records with data from 92 plant taxonomic profiles spread through diverse geographic locations all over the globe is an open domain giving a prospect for context-based scientific examination on essential oils. Our study referred to 60% trifluoro propyl methyl 40% dimethyl polysiloxane (DB-210) and 20% cyano propyl phenyl 80% dimethyl polysiloxane (DB-1701 with 36 chemicals and 130 compounds) as its chemical compounds to be perceived in large enough data sets. EssOilDB is a valuable database, for investigators to design and detection of new aroma profiles.

5 Results and Analysis

Naïve Bayes (NB), Neural Networks (NN), Multi-layer Perception (MLP), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Logistic Regression (LR) techniques are compared, and their performance metrics are given in Table 1 and Fig. 1. The accuracy (%), precision (%), sensitivity (%), f-measure (%), Matthews correlation coefficient (MCC) (%), false-positive rate (FPR) (%), false-negative rate (FNR) (%), and true negative rate (TNR) (%) values of the selected algorithms are measured (Table 1 and Fig. 1).

Our selected machine learning models have detected the retention index of gas chromatography of the essential oils. The mean square error (MSE) is measured

Table 1 Results and comparison

Model	Accuracy (%)	Precision (%)	Sensitivity (%)	F-measure (%)	MCC (%)	FPR (%)	FNR (%)	TNR (%)
NB	81	81	80	81	70	12	20	85
NN	80	79	82	79	69	11	19	83
MLP	82	84	81	81	71	12	19	87
SVM	72	70	65	62	51	13	49	82
DT	72	72	63	72	50	4	29	74
RF	82	83	71	82	66	11	21	88
LR	90	91	88	90	91	5	4	92

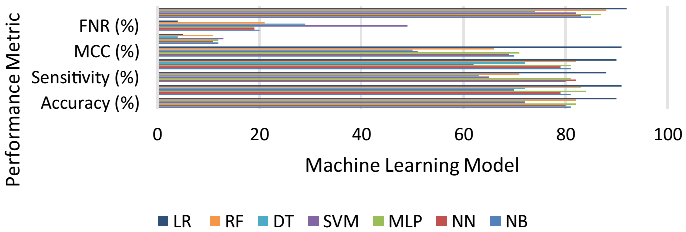


Fig. 1 Machine learning models compassion

using test datasets in the stationary phases to endorse the prediction accuracy of the selected machine learning models. The simulation outcomes suggest that the accuracy of the LR model is greater when compared with other selected methods.

6 Conclusion and Future Work

Overall, the machine learning approach can be an effective practice in producing a suitable substitute for the subjective approaches of quality assessment of essential oils. Machine learning approaches can overcome the limitations of sensory evaluation—vulnerability to large sources of deviation and the time-consuming process of it. The grouping of headspace chromatography and suitable machine learning for essential oil classification rises the speed and efficiency of the quality assessment. The present study takes the ability of various machine models into account over diverse geographical roots. In the future, we extend our work to propose an efficient machine learning model for better classification of the essential oils.

References

1. Patsilinakos A, Artini M, Papa R, Sabatino M, Božović M, Garzoli S, Vrenna G, Buzzi R, Manfredini S, Selan L, Ragno R (2019) Machine learning analyses on data including essential oil chemical composition and in vitro experimental antibiofilm activities against staphylococcus species. Molecules 24(5):890. https://doi.org/10.3390/molecules24050890
2. Siavash R, Eslam P, Mohammad RG, Parviz N (2009) Investigation of different linear and nonlinear chemometric methods for modeling of retention index of essential oil components: concerns to support vector machine. J Hazard Mater 166(2–3):853–859. ISSN 0304-3894. https://doi.org/10.1016/j.jhazmat.2008.11.097
3. Papa R, Garzoli S, Vrenna G, Sabatino M, Sapienza F, Relucenti M, Donfrancesco O, Fiscarelli EV, Artini M, Selan L, Ragno R (2020) Essential oils biofilm modulation activity, chemical and machine learning analysis—application on staphylococcus aureus isolates from cystic fibrosis patients. Int J Mol Sci 21(23):9258. https://doi.org/10.3390/ijms21239258
4. Ibrahim AF, Al-Dhaif R, Elkatatny S, Al Shehri D (2021) Machine learning applications to predict surface oil rates for high gas oil ratio reservoirs. ASME J Energy Resour Technol 144(1):013003. https://doi.org/10.1115/1.4052485
5. Ragno R, Papa R, Patsilinakos A (2020) Essential oils against bacterial isolates from cystic fibrosis patients by means of antimicrobial and unsupervised machine learning approaches. Sci Rep 10(2653):3161. https://doi.org/10.1038/s41598-020-59553-8
6. Sabatino M, Fabiani M, Božović M, Garzoli S, Antonini L, Marcocci ME, Palamara AT, De Chiara G, Ragno R (2020) Experimental data based machine learning classification models with predictive ability to select in vitro active antiviral and non-toxic essential oils. Molecules 25(10):2452. https://doi.org/10.3390/molecules25102452
7. Taghadomi-Saberi S, Mas Garcia S, Allah Masoumi A, Sadeghi M, Marco S (2018) Classification of bitter orange essential oils according to fruit ripening stage by untargeted chemical profiling and machine learning. Sensors 18(6):1922. https://doi.org/10.3390/s18061922
8. Mancianti F, Ebani VV (2020) Biological activity of essential oils. Molecules 25(3):678. https://doi.org/10.3390/molecules25030678
9. Mourad K, Ilias M, Meryem EJ, Abdelaziz B, Yvan VH (2020) Recent advances in untargeted and targeted approaches applied in herbal-extracts and essential-oils fingerprinting—a review. J Pharm Biomed Anal 177:112849. ISSN 0731-7085. https://doi.org/10.1016/j.jpba.2019.11284
10. Leo L, Alireza G, Brett P (2021) Data handling and data analysis in metabolomic studies of essential oils using GC-MS. J Chromatogr A 1640:461896. ISSN 0021-9673. https://doi.org/10.1016/j.chroma.2021.461896
11. Riela S et al (2008) Effects of solvent-free microwave extraction on the chemical composition of essential oil of *Calamintha nepeta* (L.) Savi compared with the conventional production method. J Sep Sci 31.6-7 (2008):1110–1117
12. El Asbahani A et al (2015) Essential oils: from extraction to encapsulation. Int J Pharm 483.1–2 (2015):220–243
13. Mohammadhosseini M (2016) Screening of the profiles of the essential oils from the aerial parts of Nepeta racemosa using classical and microwave-based methods: comparison with the volatiles using headspace solid-phase micro-extraction. J Herbal Drugs (An Int J Med Herbs) 7.2:97–108
14. Liang Z, Zhang P, Fang Z (2022) Modern technologies for extraction of aroma compounds from fruit peels: a review. Crit Rev Food Sci Nutr 62.5:1284–1307
15. Mwaurah PW et al (2020) Novel oil extraction technologies: Process conditions, quality parameters, and optimization. Compr Rev Food Sci Food Saf 19.1:3–20
16. Hanif MA et al (2019) Essential oils. Essential oil research. Springer, Cham, pp 3–17
17. Tohidi B, Rahimmalek M, Trindade H (2019) Review on essential oil, extracts composition, molecular and phytochemical properties of Thymus species in Iran. Ind Crops Prod 134(2019):89–99

18. Bellou E et al (2020) Laser-induced breakdown spectroscopy assisted by machine learning for olive oils classification: the effect of the experimental parameters. Spectrochim Acta Part B Atomic Spectroscopy 163:105746
19. Artini M et al (2018) Antimicrobial and antibiofilm activity and machine learning classification analysis of essential oils from different mediterranean plants against pseudomonas aeruginosa. Molecules 23.2:482
20. Lebanov L et al (2020) Random Forests machine learning applied to gas chromatography–Mass spectrometry derived average mass spectrum data sets for classification and characterisation of essential oils. Talanta 208:120471
21. Sabatino M et al (2020) Experimental data based machine learning classification models with predictive ability to select in vitro active antiviral and non-toxic essential oils. Molecules 25.10:2452
22. Patsilinakos A et al (2019) Machine learning analyses on data including essential oil chemical composition and in vitro experimental antibiofilm activities against Staphylococcus species. Molecules 24.5:890

Detecting, Analyzing, and Evaluation of Vulnerabilities Using Metasploitable

K. Thaufiqa Banu and M. Deepthi

Abstract In this paper, we will talk about penetration testing in general and dive into the specifics of using Metasploit on Metasploitable 2. It would be considered hacking to do this on any other system; therefore, I choose to use the vulnerable Metasploitable 2 instead. This study's overarching goal is to expose the wide range of methods employed in the hunt for security flaws in computer systems. System testing with Metasploit helps identify security flaws, so they can be patched, and the system is made more robust against attack. Network protocols, firewalls, and other foundational security concerns will be the focus of this study. Although there are other approaches to penetration testing, I have settled on using Metasploit due to its versatility and ease of use. We can use the highly automated community edition of the program, or we can access the Metasploit command line. The paper will investigate both of these possibilities. I will demonstrate how to efficiently use Metasploit and its associated tools to locate security flaws in a target system. Anyone who reads this document and follows the instructions can then try to exploit any system they suspect as being weak.

Keywords Stuxnet · Metasploit · Security testing · Vulnerability assessment · Exploits · Nmap · Kali Linux

1 Introduction

A total of 92% of the respondents claimed that cyber theft causes significant damage because of the loss of customer confidence and decreased corporate earnings, and in middle-sized companies, virus attacks, spyware, and backdoors were the most common forms of cybercrime. The survey was carried out by Symantec [1] and involved the investigation of 2100 businesses and government institutions from 27 countries over the course of the previous 12 months. Several different security surveys

K. Thaufiqa Banu (✉) · M. Deepthi
Department of IT, G Narayanamma Institute of Technology and Science, Hyderabad, India
e-mail: thaufiqa99@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_18

have found that viruses, penetration systems, denial of service attacks, insider abuse, spoofing, laptop theft, network/data tampering, and unauthorized insider access are some of the most serious security concerns. A total of 66% of the firm believed that the penetration system was the most serious threat, despite the fact that the virus was the most significant threat. Security in data communication networks is meant to safeguard the space that already exists inside a network or emerges as a result of a lack of experience on the part of the network administrator in the process of maintaining the security of data communication networks.

Because it is possible for someone who is not authorized to use it in order to conduct a crime, the space in the data transmission network has the potential to be lethal. There are many companies that offer security services, but the majority of the time, these companies do not associate themselves with the appropriate security systems. This section of the article will walk you through each stage of scanning the vulnerabilities and backdoors connected with the platforms. The purpose of the technique known as "penetration testing" is to protect the systems from any potentially destructive attacks. In WHOE, the procedure assists in understanding the security integrity and the kind of attacks that were carried out by the attacker, and it also helps to grant authorization to prevent such things from happening in the future. Penetration testing is carried out on the client's authorized system, and the goal is to locate any backdoors in the system so that they may be patched and made more secure.

2 Related Work

Using the standards laid out by Kitchenham in his SLR recommendations, Ah mad Salah Al-Ahmad et al. have conducted a comprehensive analysis of existing penetration testing models and procedures. According to the SLR findings, there are a few places where more work needs to be done: developing a model for MCC application penetration testing; conducting research on mobile, cloud, and web vulnerabilities; and developing research on these topics. In order to expose previously unknown security flaws, promote trust among participants, and give programmers the tools they need to create more secure MCC apps, two new and crucial needs must be met: offloading and mobile state management. Positive findings from the review that can inform subsequent studies are presented.

According to Prashant Vast et al., a PEN tester can examine the system's defense mechanisms and determine how vulnerable it is to network security and intrusion assaults by conducting a penetration test. In this study, we analyzed the previous research on penetration (PEN) testing through a literature review. We have made an effort to cover a lot of ground with this overview of PEN testing. Furthermore, we have researched the numerous PEN testing tools in terms of their usefulness, technical specs, release date, platform compatibility, etc.

Michele Fiocca's study of penetration testing will look into questions including how easy it is to hack into a system when access is limited or the target is located far away, as well as how much work is involved in doing so. It is possible to get unauthorized access to a target system through a process known as "pen testing," which employs a variety of tools. Thus, a penetration tester's primary responsibility is to identify security holes in new software and infrastructure and alert the development team, so that they may make the necessary adjustments. There is a discussion of the many tools employed during the various penetration testing phases included in this report.

For those unfamiliar with penetration testing, Ken van Wyk's piece gives a useful primer. After introducing the challenges of incorporating penetration testing into a software development life cycle, the article describes the drawbacks of conventional penetration testing methods and offers suggestions for their improvement. The different kinds of penetration testing tools and some examples may be found in this connected article.

According to the scenario described by Justin Pierce et al., in which commercial software is constantly patched to fix security problems, penetration testing can give businesses an accurate picture of their security. By mimicking the methods employed by malicious hackers, penetration testing is able to confirm the absence or presence of security measures in a network.

3 Research Method

The system being used in the study is Metasploitable 2. It is an OS based on Linux, but it was developed specifically with Metasploit in mind. If you want to undertake penetration testing, you may get it for free and quickly through the Metasploit website. Software security holes can be understood via the lens of the study's methodology, which can be broken down into the following steps. The findings of a penetration test can help administrators strengthen their network's defenses. However, there are a number of moral concerns that arise when hackers are involved in the testing process, all of which revolve around the tester's responsibility to uphold professional standards. In this study, we provide our conceptual model and improved taxonomy, and we examine the ethics of penetration testing. Penetration testing, computer security, and computer ethics are some examples of keywords. To get started, grab the vulnerable machine image from www.metasploit.com, then load it up in your preferred virtualization software. By now you should be well on your way to becoming a penetration tester and testing the system's weaknesses. After that, you can quickly join by entering msfadmin as the username and leaving the password field blank. Despite the fact that it is merely a test system, it contains all the features of any operating system that we could want to test in the future (Fig. 1).

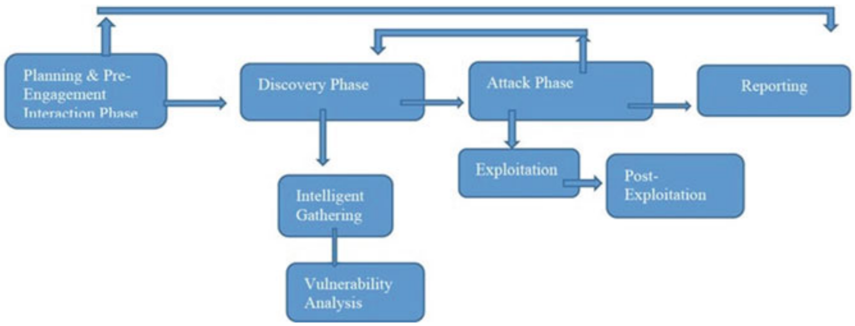


Fig. 1 Design flow method

3.1 Metasploit Framework

The most popular penetration testing framework is called Metasploit. It was first conceived of by its creator, H.D. Moore, in 2003. Rapid7, a private corporation that makes several cyber security tools, bought the Metasploit project in 2009. The source code for the Metasploit framework is available on GitHub and is licensed under the BSD-3-clause open-source license. While Perl was used to creating the first version, and Ruby was quickly adopted as the language of choice for platform development. The msfconsole and msfcli applications are the recommended means of interfacing with the Metasploit framework on Linux-based systems, especially headless ones (without peripherals, accessible solely over SSH, or a similar protocol). From host discovery to obtaining a root shell, msfconsole provides a fully interactive session for conducting a penetration test. The following is a synopsis of the exploitation phase of Metasploit penetration testing:

- Use a host discovery tool, such as Nmap or Metasploit's discovery modules.
- Compile a list of prospective target services by running vulnerability scanning tools like Nessus, Open-Vas, or Nexpose against (a subset of) the discovered hosts.
- Select an appropriate attack for a vulnerable service identified in Step 2. The exploits are built by a group of penetration testers, who are often the same persons that filed a CVE entry for the flaw in question.
- Decide what to put on the infected machine (the "payload"). The meterpreter shell, which allows for the execution of arbitrary commands on the target, is the payload with the greatest features. It conceals its presence on the disc by injecting itself as a dynamically loaded library (DLL) into a running program. For added security, the meterpreter establishes a Transport Layer Security v1 (TLSv1) session for exchanging data and installing extensions [19].
- Attempting to trick IDS by encoding the attack flow.
- Launching an attack on the target using the exploit, payload, and encoder combination.

The modular design of Metasploit made it the most widely used tool, as it allowed for the combination of any exploit with any payload and any encoding. Currently, Rapid7's exploit-DB [20] contains information on 3664 different exploits. The professional edition of the Metasploit framework includes all of these features and more. Penetration testing is a time-consuming and routine activity, i.e., because the attacker needs to work in tandem with his weapons. It is possible to exploit a service if the configuration is incorrect, but not with the default settings. Attacker expertise is still crucial, and therefore giving him access to his equipment is essential. Some work, however, is amenable to automation, particularly data collection.

4 Result and Discussion

The fundamentals of console navigation for manual system exploitation. An exploit for a specific system, and not necessarily Metasploitable 2, will be demonstrated in this method.

The first thing to do is to launch the console and log in as msfadmin (Fig. 2).

The next step is to choose the target system and the method of exploitation you intend to use.

After you have found that out and received the system's IP address, you can input Nmap followed by the IP address. To determine which ports are accessible, the console creates a map of the system's ports (Fig. 3).

- Once we know the port is open, we can use display commands to look for vulnerabilities, and we identify one that targets a remote computer.
- The command 'info exploit/windows/dcerpc/ms03 026 dcom' will provide additional details on the exploit.
- Next type on terminal use "exploit" into the console.

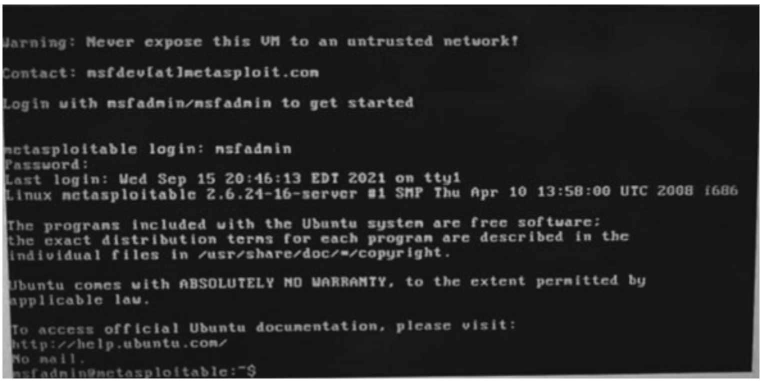


Fig. 2 Launch console and login as msfadmin

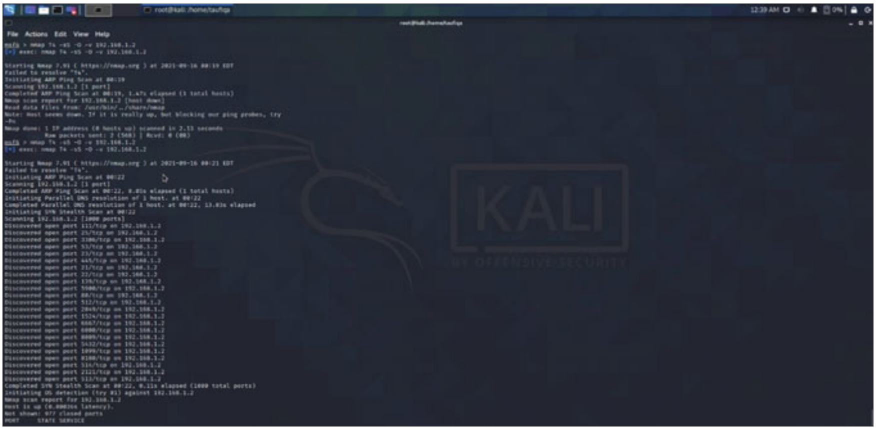


Fig. 3 Determine accessible ports

- Once that is loaded, type "display options" to see what to do next in order to advance the exploit.
- After that, we need to discover the payload to push into the open port and provide the IP address that we want to exploit (set RHOST "ip address").
- If we want to locate payloads that work with this vulnerability, we need to type show payloads into the input field. After that, type "set LHOST 192.168.1.2" and "set PAYLOAD windows/meterpreter/reverse tcp" into the console (Figs. 4, 5 and 6).

Now, type the check command to determine whether or not the "computer is vulnerable to the exploit or not," and if so, proceed with the command, exploit, to carry it out (Fig. 7).

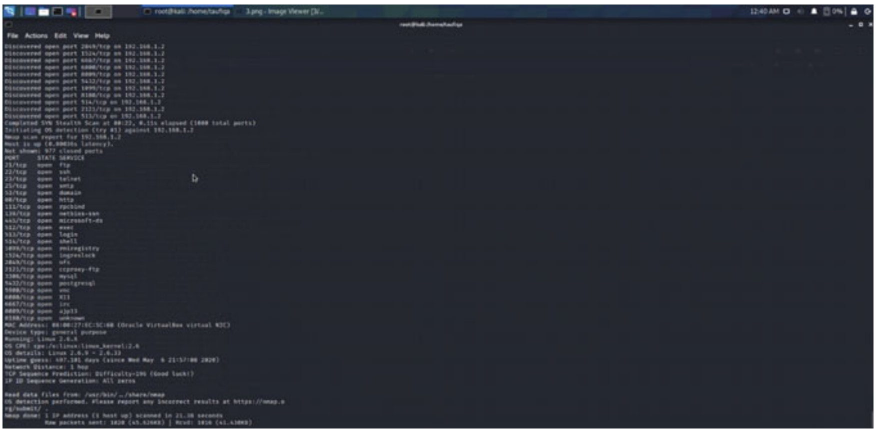


Fig. 4 Display commands to vulnerabilities

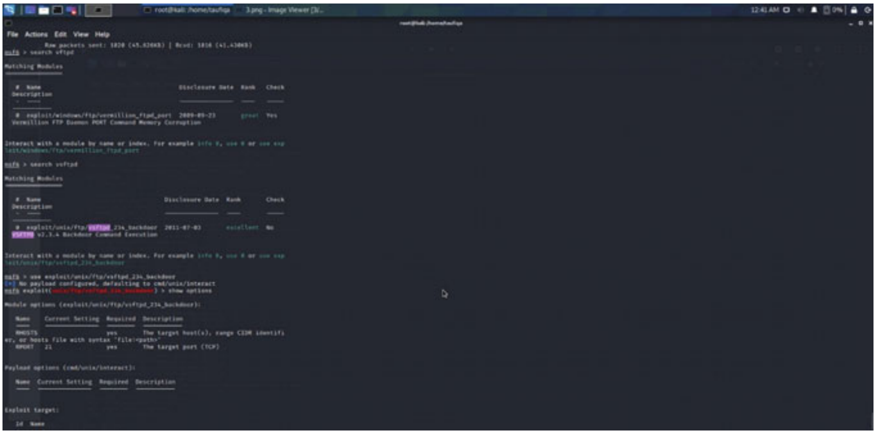


Fig. 5 Info exploit

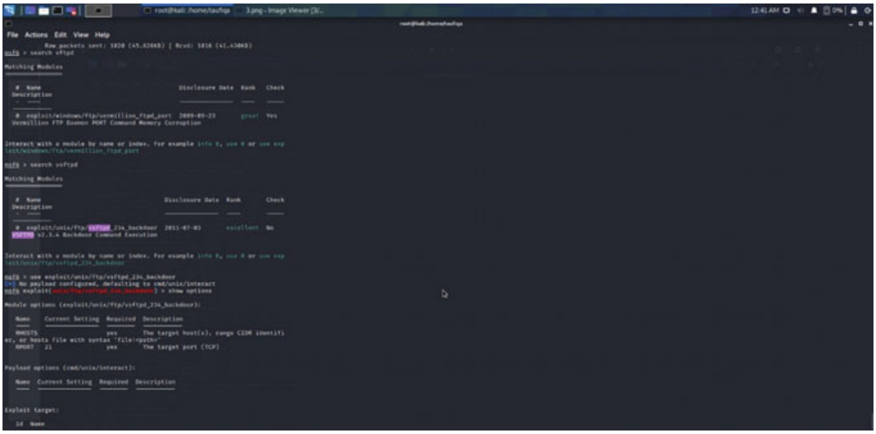


Fig. 6 Exploit

Access to the system and complete freedom to perform any desired actions within the system should be granted if all goes according to plan.

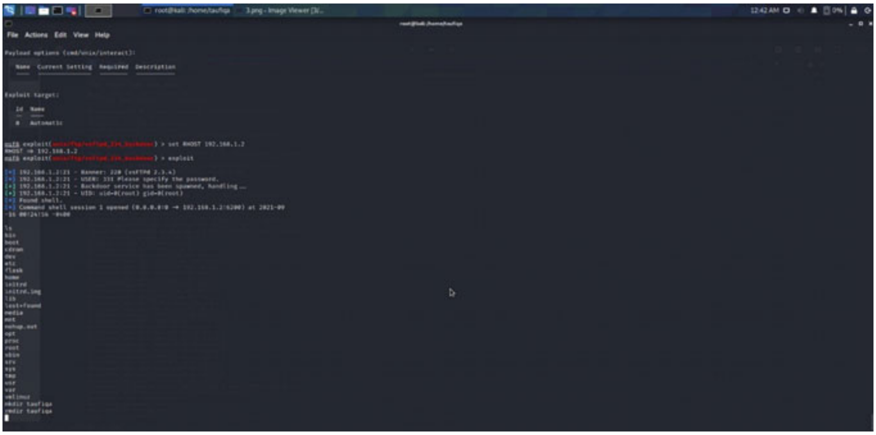


Fig. 7 Discover the payload

5 Conclusion

Many different tools exist for performing penetration tests, but Metasploit is the most useful one in my experience and is the one I will recommend to you. You may use it manually or automatically, and it comes with a lot of cool features. In spite of the fact that I have already demonstrated the arguments for and against the two in this paper, I would want to restate a couple of my original points. Manually exploiting a system gives you more flexibility over how you try to exploit it, but it also takes longer. If you want to be sure your data is safe from hackers, one method is to do penetration tests on your systems. If you want to undertake penetration testing, I promise you will not regret trying out Metasploit. There is a plethora of software out there for online use. Caution is required while using any of the applications because mistakes can have disastrous consequences.

References

1. Denis M, Zena C, Hayajneh T (2016) Penetration testing: concepts, attack methods, and defense strategies. In: 2016 IEEE long Island systems, applications, and technology conference (LISAT)
2. Stefinko Y, Piskozub A, Banakh R (2016) Manual and automated penetration testing. Benefits and drawbacks. In: 13th international conference on modern problems of radio engineering, telecommunications and computer science
3. Gupta H, Kumar R (2015) Protection against penetration attacks using Metasploit. In: 2015 4th international conference on reliability, Infocom technologies and optimization (ICRITO) (Trends and Future Directions)
4. Talekar N, Penetration testing with Metasploit framework | www.SecurityXploded.com, SecurityXploded.com
5. Kennedy D (2011) Metasploit: the penetration tester's guide. No Starch Press, San Francisco, CA

6. Hack O, Like a pro-null Byte ń wonder how to, WonderHowTo. https://null-byte.wonderhowto.com/how-to/hack-like-a-pro/
7. Holik F, Horalek J, Marik O, Neradova S, Zitta S (2014) Effective penetration testing with Metasploit framework and methodologies. In: 2014 IEEE 15th international symposium on computational intelligence and informatics (CINTI)
8. Finding IP Addresses of Other Network Interfaces on Linux, pentest monkey. http://pentestmonkey.net/finding-ip-addresses-of-other-network-interfaces
9. How to use Metasploit commands for real-world security tests, SearchSecurity. http://searchsecurity.com/tip/Using-Metasploit-for-real-worldsecurity-tests
10. Dodd D, Penetration testing and shell tossing with Meta… ż ADMIN magazine, ADMIN Magazine. http://www.admin-magazine.com/Articles/Pen-Test-Tips
11. Hack the Fartknocker VM (CTF Challenge), Hacking Articles, 06-Apr-2017. http://www.hackingarticles.in/hack-fartknocker-vm-ctf-challenge/
12. Masood R, U-E-G, Anwar Z (2011) SWAM: Stuxnet worm analysis in Metasploit, 2011 frontiers of information technology
13. Fossi M, Turner D, Johnson E, Mack T, Adams T, Blackbird J et al (2010) Symantec internet security threat report-trends for 2009, XV(April)
14. Stallings W (2005) Cryptography and network security: principles and practices. Crypt Netw Secur 592
15. Presekal A, Sari RF (2014) Performance comparison of host identity protocol and TCP/IP with firewall against denial of services. Indonesian J Electr Eng Comput Sci 12(12):83358343
16. Kahtan Mohammed R, Yoichiro U (2018) An FPGA-based network firewall with expandable rule description. Indonesian J Electr Eng Comput Sci 10(3):13101318
17. Stuart D, Beaver K (2013) Next-generation IPS For Dummieső. Sourcefire edition. Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki:69
18. Sekhar R, Perumal K, Rani SV (2015) Analysis of next generation intrusion prevention system using sensor fusion and fuzzy logic. Int J Sci Res Eng Technol (IJSRET) 4(9):9368
19. Kamara S, Fahmy S, Schultz E, Kerschbaum F, Frantzen M (2003) Analysis of vulnerabilities in internet firewalls. Comput Secur 22(3):21432
20. Ho C-Y, Lin Y-D, Lai Y-C, Chen I-W, Wang F-Y, Tai W-H (2012) False positives and negatives from real traffic with intrusion detection/prevention systems. Int J Futur Comput Commun. 1(2):87
21. Ho CY, Lai YC, Chen IW, Wang FY, Tai WH (2012) Statistical analysis of false positives and false negatives from real traffic with intrusion detection/prevention systems. IEEE Commun Mag 50(3):14654
22. Stiawan D, Abdullah AH, Idris MY (2011) Characterizing network intrusion prevention system. Int J Comput Appl 14(1):9758887
23. Sourcefire (2014) Sourcefire 3D system user guide. Simulation 1280
24. Cisco System. FireSIGHT System user guide creating traffic profiles. Simulation
25. Software Testing (2017) Penetration testing—complete guide with penetration testing sample test cases software testing help
26. Lakshmi DR, Mallika SS (2017) A review on web application testing and its current research directions. Int J Electr Comput Eng
27. Clarke J, Alvarez RM (2009) SQL injection attacks and defense:494
28. Shinde G, Waghere SS (2017) Analysis of SQL injection using DVWA tool
29. Cannady JD (2000) Next generation intrusion detection: autonomous reinforcement learning of network attacks. Proc 23rd Natl Inf Syst Secure Conf 112
30. Boonen R (2014) Fuzzy security game over Dolo Malo-JavaScript adware exposed

An Enhanced Machine Learning Technique to Predict Heart Disease

K. Shilpa and T. Adilakshmi

Abstract Cardiovascular heart diseases are most important demise of life. The early prediction of cardiovascular disease of heart is challenging task in the clinical healthcare. Machine learning is part of artificial intelligence. In this, we have considered some feature which can lead to illness of heart. An enhanced technique of machine learning to predict heart disease using different features has been proposed. The aim is to detect best classification algorithm for disease prediction with maximum accuracy. We have taken datasets from UCI ML dataset with 1025 instances and 14 attributes. We have applied three different ML algorithms such as, Naïve Bayes, K-nearest neighbor, and decision tree. We proposed decision tree technique with 100% accuracy when compared to Naïve Bayes and K-NN, and moreover, time taken to build model is less when compared to other algorithm. The second place is taken by Naïve Bayes algorithm with 86.38% accuracy. Third place is taken by K-NN which has given 85.99% accuracy. This result will benefit to select the best classification algorithm for heart disease prediction and can be used for detection and treatment.

Keywords Machine learning · UCI ML dataset · Naïve Bayes · K-NN · Decision tree · Python

1 Introduction

Machine learning is branch of AI, which uses many algorithms to extract patterns from raw data. ML is to permit computer systems to learn out of experience aside from human intervention or explicitly programmed [1–5].

K. Shilpa (✉)
Research Scholar, Department of CSE, Osmania University, Hyderabad, India
e-mail: shilpamtech555@gmail.com

Assistant Professor, Department of CSE, CMR Technical Campus, Hyderabad, India

T. Adilakshmi
Professor & Head, Department of CSE, Vasavi College of Engineering, Osmania University, Hyderabad, India
e-mail: t_adilakshmi@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_19

Heart disease or cardiovascular diseases (CVD) are mostly leading to demise or death. CVD can attack because of family health history and unhealthy habits like smoking, alcohol, obesity, physically unfit, and imbalanced diet. Now early detection of CVD has to done to increase survival rate.

The heart attack symptoms are the arm of deadness, face or leg or body of one side, unable see with both eyes, unconscious or fainting, unable to walk, and severe headache. Cardiovascular disease (CVD) types are coronary HD, H-stroke, arterial disease, and aortic disease.

1.1 Objectives and Motivation

The objective is to identify the best machine learning algorithm for prediction of the heart disease and toward finding accuracy and performance of different machine learning algorithm and differentiate precision, accuracy, recall, and F1 Score.

To increase the survival rate when early detection of prediction of heart disease and right diagnosis leads to right treatment.

2 Literature Work

Singh and Kumar [6] have predicted heart disease using various ML algorithms such as, SVM, k-nearest neighbor, LR, and DT with accuracy. K-NN has given highest accuracy 87% compared to SVM which has given 83%, decision tree has given 79%, and linear regression has given 78%.

Gavhane et al. [7] have built an application of disease of heart that can predict disease of heart based on various attributes such as pulse rate, BP, HRate, diabetes, and hyper cholesterol. They have used some ML algorithms like multi-layer perception (mlp) and neural network.mlp as proposed system, and it has improved efficiency and accuracy.

Nikam et al. [8] have used different feature for predicting cardiovascular disease. They used vital feature BMI which is used to predict heart disease. The main objective is to find effect of BMI on the heart disease prediction. Apart from other feature, adding BMI has improved the accuracy of ML algorithms. They have used DT and XGB classifiers among this decision tree which has given highest accuracy.

Salhi et al. [9] have done pre-processing and selected relevant feature, and then they have applied three techniques such as NN, K-NN, and SVM with different sizes. Among three techniques, NN has achieved 93% of accuracy.

Mienyea et al. [10] have used cart technique, and it generates homogeneous ensemble from various models of cart based on weighted aging classifier ensemble. It has given better performance. They have applied techniques on two datasets such as Cleveland and Framingham datasets which have got accuracies 93% and 91%, respectively.

3 Methodology

Decision tree is a tree-like structure, each internal signifies the attribute/feature, and decision rules signifies by branches and outcome/class by leaf node. Gini index and information gain are attribute measures.

$$\text{Information Gain} = \text{Entropy}(\text{Parent}) - [\text{Average Entropy}(\text{Children})]$$

Gini Index: To construct the DT and function that determines how well a DT was split

$$\text{Gini Index} = 1 - \Sigma_i P_i^2$$

Naïve Bayes

Naïve Bayes is collection of classification algorithms, and it depends on Bayes theorem which is based on conditional probability.

K-Nearest Neighbor

K-NN is simplest algorithm of machine learning based on supervised technique. It selects the number of K of the different neighbors. The K no of neighbors is based on Euclidean distance.

3.1 Proposed Work Flow

See Fig. 1.

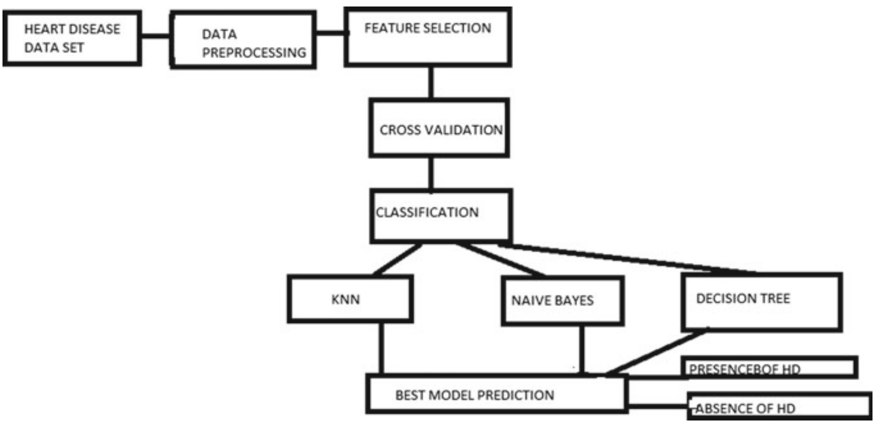


Fig. 1 Proposed work flow of heart disease prediction

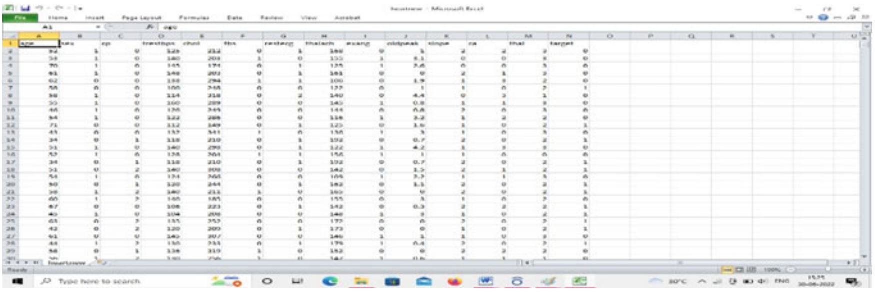


Fig. 2 Heart disease prediction dataset

3.2 *Heart Disease Prediction Data Set*

Data has been collected from machine learning repository of UCI with 1025 instances and 14 attributes/feature where target indicates disease prediction like 0 indicate NO disease and 1 indicates yes disease (Fig. 2).

4 Results and Discussion

We have applied three ML algorithms such as Naïve Bayes, K-NN, and decision tree on dataset. The following are the calculations

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{1}$$

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{2}$$

$$\text{Recall} = \frac{T_p}{T_p + T_n} \tag{3}$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

See Fig. 3; Tables 1, 2, 3 and 4.

Experimental Tool

We have used Python 3.6.8 anaconda spider (Figs. 4, 5 and 6).

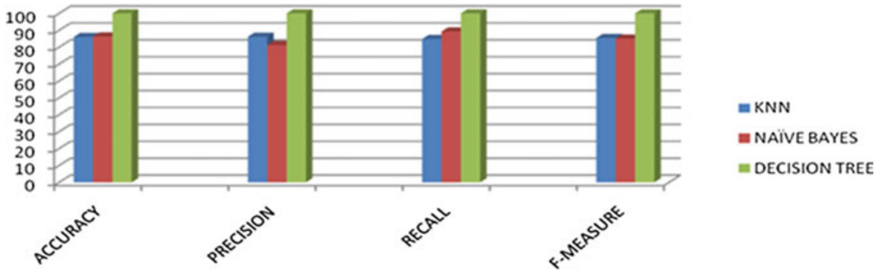


Fig. 3 Precision, accuracy, recall, and F-measure of different ML algorithms

Table 1 Precision, accuracy, recall, and F-measure of DT

Confusion Matrix of Decision Tree

123	
0	

ACCURACY	$\frac{123+134}{123+134+0+0}$	100%
PRECISION	$\frac{123}{123+0}$	100%
RECALL	$\frac{123}{123+0}$	100%
F-MEASURE	$\frac{123}{123+1/2(0+0)}$	100%

Table 2 Precision, accuracy, recall, and F-measure of NB

Confusion Matrix of Naïve Bayes

100	23
12	122

ACCURACY	$\frac{100+122}{100+122+12+23}$	86.38%
PRECISION	$\frac{100}{100+23}$	81.30%
RECALL	$\frac{100}{100+12}$	89.28%
F-MEASURE	$\frac{100}{100+1/2(23+12)}$	85.106%

Table 3 Precision, accuracy, recall, and F-measure of K-NN

Confusion Matrix of K-NN

106	17
19	115

ACCURACY	$\frac{106+115}{106+105+17+19}$	85.99%
PRECISION	$\frac{106}{106+17}$	86.17%
RECALL	$\frac{106}{106+19}$	84.80%
F-MEASURE	$\frac{106}{106+1/2(17+19)}$	85.48%

Table 4 Precision, accuracy, recall, and F-measure of different ML algorithms

	Naïve Bayes (%)	K-NN (%)	Decision tree (%)
Accuracy	86.38	85.99	100
Precision	81.3	86.17	100
Recall	89.28	84.8	100
F-measure	85.11	85.47	100

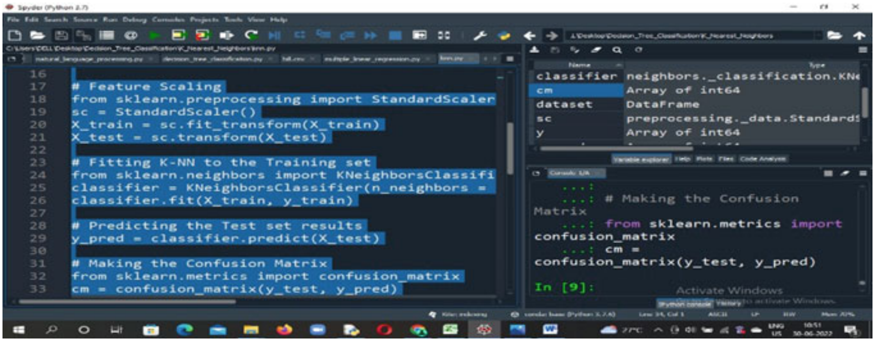


Fig. 4 K-NN

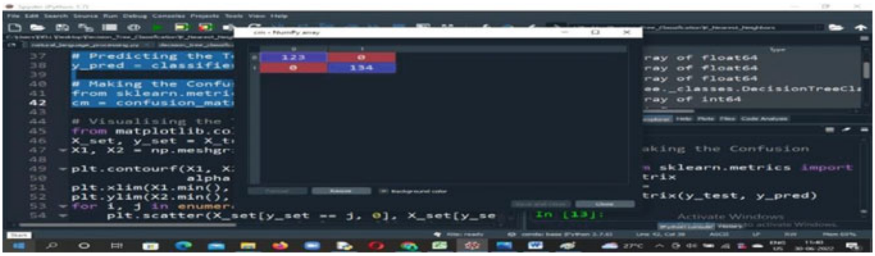


Fig. 5 Confusion matrix of decision tree

Fig. 6 Confusion matrix of Naïve Bayes

5 Conclusion and Future Work

We proposed an enhanced technique of machine learning to predict heart disease using different features. The key goal is to detect best classification algorithm for disease prediction with maximum accuracy. We have taken datasets from UCI ML dataset with 1025 instances and 14 attributes. We have applied three different ML algorithms such as Naïve Bayes, K-NN, and decision tree. We proposed decision tree technique with 100% accuracy when compared to Naïve Bayes and K-NN, and moreover time taken to build model is less when compared to other algorithm. The second place is taken by Naïve Bayes algorithm with 86.38% accuracy. Third place is taken by K-NN which has given 85.99% accuracy. This result will benefit to select the best classification algorithm applied for detection and treatment. In future, we want to implement with real-time dataset with deep learning techniques to improve the accuracy.

References

1. Mohan SK, Thirumalai C, Srivastva G (2019) Effective heart disease prediction using hybrid machine learning techniques. IEEE Access
2. Sharma H, Rizvi MA (2017) Prediction of heart disease using machine learning algorithms: a survey. Int J Recent Innov Trends Comput Commun 5(8)
3. Nikhil Kumar M, Koushik KVS, Deepak K (2019) Prediction of heart diseases using data mining and machine learning algorithms and tools. Int J Sci Res Comput Sci Eng Inf Technol
4. Kaur A, Arora J, Heart diseases prediction using data mining techniques: a survey. Int J Adv Res Comput Sci, IJARCS 2015–2019
5. Kohli PS, Arora S (2018) Application of machine learning in diseases prediction. In: 4th international conference on computing communication and automation (ICCCA)
6. Singh A, Kumar R, Heart disease prediction using machine learning algorithms
7. Gavhane A, Kokkula G, Panday I, Devadkar K (2018) Prediction of heart disease using machine learning. In: Proceedings of the 2nd international conference on electronics, communication and aerospace technology (ICECA)
8. Nikam A, Bhandari S, Mhaske A, Mantri S, Cardiovascular disease prediction using machine learning models
9. Salhi DE, Tari AK, Tahar Kechadi M, Using machine learning for heart disease prediction
10. Mienyea ID, Sun Y, Wang Z, An improved ensemble learning approach for the prediction of heart disease risk

3D Vision Transformer for Cervical Spine Fracture Detection and Classification

Satwik Sunnam, Avuku Obulesh, Sri Charan Mohan Janthuka, and Keerthi Yalamaddi

Abstract In USA alone, there are over 1.5 million spine fractures per year, resulting in about 17,730 spinal cord injuries. The cervical spine is where fractures in the spine most frequently occur. The prevalence of spinal fractures in the elderly has increased, and in this population, fractures may be harder to see on imaging because of coexisting degenerative illness and osteoporosis. Nowadays, computed tomography (CT) is almost completely used instead of radiography for the imaging diagnosis of adult spine fractures (X-rays). To stop neurologic degeneration and paralysis following trauma, it is vital to trace any vertebral fractures at the earliest. Many approaches have been proposed for the classification of the cervical spine [2d models]. We are here in this paper trying to break the bounds and use the vision transformers, a SOTA model in image classification, by making minimal changes possible to the architecture of ViT and making it 3D-enabled architecture, and this is evaluated using a weighted multi-label logarithmic loss.

Keywords I cervical spine · Spinal fractures · Osteoporosis · Computed tomography (CT) · 2D models · ViT · Multi-label logarithmic loss · Kaggle · Public score · Private score

1 Introduction

Our spinal column's neck area contains the cervical spine. These are the first seven bones in your body (C1–C7). Your cervical spine is surrounded by intervertebral discs, the spinal cord, nerves, muscles, tendons, and ligaments, among other things. A cervical fracture is damage that occurs to the bones which may result from compression (minor trauma in a person with osteoporosis), a burst fracture (vertebra that is

S. Sunnam (✉) · A. Obulesh · S. C. M. Janthuka · K. Yalamaddi
Department of Artificial Intelligence, Vidya Jyothi Institute of Technology (Autonomous), Hyderabad, India
e-mail: sunnamsatwik19@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_20

crashed in all directions), or a fracture-dislocation (mostly from a vehicle accident or falling from heights).

Advancement of the technologies in the domain of AI helps many medical surgeries. Trials are being held whether robotic surgeries or smart software can detect missing information or location by doctors during surgeries. Improvements in deep learning usage helped other authors explore the possible methodologies for detecting cervical spine fractures. A CNN-based model ResNet-50 [1] has been utilized with the Bi-LSTM utilizing the features from CNN, and remembrance of Bi-LSTM was utilized to build a mechanism for detection [2]; the results are evaluated on various metrics such as F1 score, Matthews correlation coefficient (MCC), and area under curve (AUC), the number of Bi-LSTM layers have been sequentially increased to see the performance, and the results seem promising.

Another paper [3] used the FDA-approved CNN developed by Aidoc to detect cervical spine fracture detection over 665 examinations, and they have also taken into consideration of various metrics such as accuracy, sensitivity, and specificity which are 92, 76, and 97%. And they have cross-checked the results with the results from the radiologists 95, 93, and 96% which represent the same sequence of metrics as the model.

A 3D U-Net is used by this paper [4], where they have considered the voxel-level representations to dig the deeper part than 2D models and minimal changes in the U-Net by replacing its upsampling method with a classification label.

Major state-of-the-art models in NLP use the transformers [5] architecture as the base for building cherries on top for down-streaming tasks. The attention mechanism [5] basically tells the network to learn important words which majorly contribute to the context of the sentence, and many papers have explored a variety of hybrid models using this mechanism and architectures [6–8], but also changing the internal workflow in attention mechanism [9, 10] helped some tasks to achieve better results than the traditional mechanism which was described in [5].

Without limiting transformers and attention mechanisms not only to NLP vision tasks which can also be done by using the transformers architecture rather than taking the entire architecture, only the encoder part is taken and the image passing to the model is broken down into patches and connecting those patches with serial numbers to represent the sequence of images (embedding); this has outperformed all the existing SOTA models on classifications tasks, and this is called as vision transformer (ViT) [11], and few papers have also explored the 3D version of this but modifying the architecture [12, 13] of the model for the particular use case.

In this paper, we are proposing the VIT 3D model with minimal changes when compared to the original work [7].

The article's organization is as follows. Section 2 describes the dataset collection and description method steps. Section 3 discusses data preparation. Section 4 discusses the methodology and related concepts. Section 5 describes the performance evaluation of the proposed methodology, and Sect. 6 gives conclusions of this article.

2 About the Dataset

The dataset was produced by the American Society of Neuroradiology (ASNR) and the American Society of Spine Radiology (ASSR). The challenge planning task force collected imaging data from twelve sites on six continents, counting approximately 3000 CT scans, to establish a ground truth dataset. Experts in spine radiology annotated these studies to specify the occurrence, vertebral level, and location of any cervical spine fractures. A subset of the imaging datasets was automatically segmented using a 3D UNET model, and the segmentations were adjusted and approved by radiologists. For C1 through C7, the given segmentation labels range from 1 to 7 (Seven cervical vertebrae).

2.1 Dataset Description

The dataset is structured as below; it consists of CSV files and image folders corresponding to the train and test data. Train.csv have metadata for the train image data.

- StudyInstanceUID—The study ID. This is the one unique study ID for each patient scan.
- Patient_overall—One of the target columns. The patient-level outcome, i.e., if any of the vertebrae are fractured.
- C [1–7]—The other target columns. Whether the given vertebrae are fractured. [7 cervical vertebrae are represented using the notation of C[I], I ranges from (1–7)].

Test.csv have metadata for the test set prediction structure.

- row_id—The row ID. This will match the same column in the sample submission file.
- StudyInstanceUID—The study ID.
- Prediction_type—Which one of the eight target columns needs a prediction in this row.

[train/test]_images/[StudyInstanceUID]/[slice_number].dcm

The image data is organized with one folder per scan. Each image is in the DICOM file format. The DICOM image files are <= 1 mm in thickness, axial, orientation, and bone kernel. Some of the images are compressed to JPEG.

sample_submission.csv A valid sample submission.

- row_id—The row ID. See the test.csv for what prediction needs to be filled in that row.
- fractured—The target column.

train_bounding_boxes.csv Bounding boxes for a subset of the training set

Fig. 1 Data preparation

3 Dataset Preparation

As we have seen in the dataset description, every patient has multiple CT scans. Meaning that we can use multiple images to predict if there are one or multiple fractures involved in the study. The preparation has three steps.

Step 1: Select the constant number [*N*] of scans from each study and sort them out using the file names [Ex: 1.dcm, 2.dcm, 3.dcm, …, N.dcm]. Each original image has a shape of [512, 512, 1].

Step 2: We stack images in the correspondence of sorting order that has a size of [384,384,*N*,1] where "*N*" represents the total number of stacked images.

Step 3: We transpose the axis as needed for the model into the desired shape (3,0,1,2).

The visual representation of the above steps is shown below (Fig. 1).

4 Methodology

Every architecture starts with basic building blocks which can start with basic building blocks which complete the entire model in further sub-sections which we will be discussing these and finally the methodology.

4.1 Encoder Mechanism

The encoder in the transformer consists of multiple encoder blocks. An input image embedding goes through the encoder blocks, and the output of the last encoder block becomes the input features to the classification head.

The internal working of the encoder is discussed below.

The encoder block employs the self-attention method to add context from the entire phrase to each token (embedding vector).

Each token may have more than one connotation and/or function, depending on the nearby tokens.

So that the model may access various embedding subspaces, and the self-attention method uses multiple heads (eight parallel attention calculations).

ReLU and another linear layer, which processes each embedding vector separately with the same weights, are both parts of the position-wise feed-forward network (FFN).

Therefore, for further transformation, each embedding vector (including contextual information from the multi-head attention) passes through the position-wise feed-forward layer.

Let x be an embedding vector. The position-wise feed-forward network is:

$$\mathrm{FFN}(x) = \mathrm{ReLU}(xw_1 + b_1)W_2 + b_2 \tag{1}$$

w_1 raises the dimension of x from 512 to 2048, and w_2 decreases it from 2048 to 512. All places inside the same layer have the same weights in FFN.

Encoder block uses three residual connections, which is simply an element-wise addition:

$$x + \mathrm{Sublayer}(x) \tag{2}$$

The prior embeddings are carried over to the following levels through residual connections.

As a result, the encoder blocks contribute more information to the embedding vectors using position-wise feed-forward networks and multi-head self-attention computations.

After each residual connection, there is a layer normalization:

$$\mathrm{LayerNorm}(x + \mathrm{Sublayer}(x)) \tag{3}$$

Layer normalization seeks to reduce the impact of covariant shift, just like batch normalization does.

In other words, it stops training from being stable and fast (i.e., we cannot make the learning rate big enough) by preventing the mean and standard deviation of embedding vector elements from moving about.

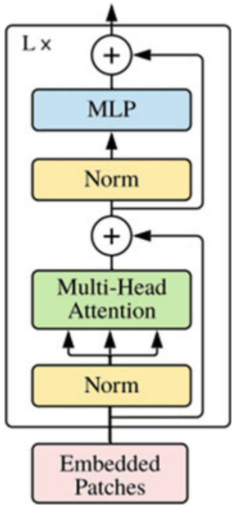
Layer normalization operates at each embedding vector, as opposed to batch normalizing (not at the batch level) (Fig. 2).

4.2 Self-attention Mechanism

The encoder–decoder network's creators suggested drawing more attention to it. This restriction, preventing the input sequence from being encoded into more than one fixed-length vector, is proposed to be lifted. It is anticipated that this issue would crop up more frequently while decoding lengthy sequences.

Unlike the attention mechanism, which allows output to focus attention on input while creating output, the self-attention model allows inputs to communicate with one another (i.e., compute the attention of all other inputs WRT one input).

Fig. 2 Transformer encoder



There are multiple steps are involved in the attention mechanism.

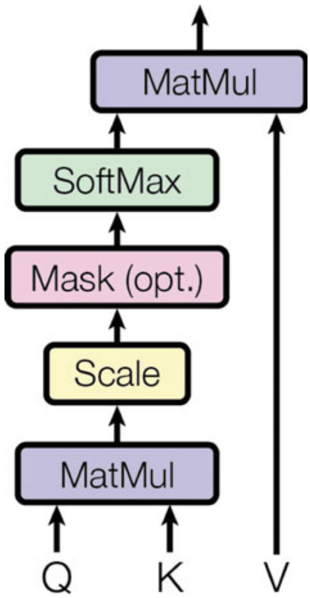
1. The initial action is the three weight matrices ($W(Q)$, $W(K)$, and $W(V)$) obtained during the first stage's training procedure multiplied by each encoder input vector. As an outcome of this matrix multiplication, we will have three vectors for all of the input vectors: the key vector, the query vector, and the value vector.
2. The key vectors from other inputs are multiplied by the query vector from the given inputs in order to calculate self-attention in the second step.
3. In the third phase, we will multiply the result by the square root of the magnitudes of the key vector (d_k). The magnitude of the key vector in the paper is 64, yielding 8 as the outcome. The reason for this is that when the dot products get huge, certain self-attention scores become extremely low in the forthcoming when the softmax function is utilized.
4. Implementing the softmax function to all of the self-attention ratings we calculated with respect to the search keyword is the fourth stage (here first word).
5. In the fifth step, the value vector is multiplied by the calculated vector from the earlier step.
6. The weighted value vectors obtained in the previous step are added to obtain the self-attention output for the given word in the final step.

All input classifications are put through the above-mentioned procedure. The self-attention matrix for the input matrices (Q, K, and V) is obtained mathematically as follows:

$$\text{Attention}\ (Q, K, V) = \text{softmax}\left(Q_K^T / \sqrt{d_k}\right) V \quad (4)$$

where Q, K, and V are the query, key, and value vectors combined (Fig. 3).

Fig. 3 Self-attention



4.3 Multi-head Attention

The researchers of the attention study developed a distinct kind of attention mechanism called multi-headed attention. The transformer's attention module executes its computations frequently and simultaneously. Each of these is known as an attention head. The attention segment ruptures its query, key, and value arguments in N distinct ways, individually routing each rupture over a different head. The result of integrating all of these linked attention computations is a final attention score. The transformer is better able to scramble different links and subtle characteristics for each word by using a method called "multi-head attention."

The multi-head attention has five steps, and they are described below:

1. Start by creating an embedding of each word in the original sentence.
2. For this mechanism, we created h ($h = N$) and N different attention heads, each with a different weight matrix [$W(Q)$, $W(K)$, or $W(V)$].
3. The key, value, and query matrices for each attention head are produced in this phase by multiplying the input matrix with weight matrices (W^Q, W^K, and W^V).
4. From each attention head, an output matrix is obtained as the attention method is executed on the query, key, and value matrices.
5. To create the outcome of the multi-headed attention layer, we mix the output matrices generated by each attention head, the dot product, and the weight W_0.

Mathematically, multi-headed attention can be denoted as:

$$\text{Multi-head}(Q, K, V) = \text{concat}(\text{head1}, \text{head2}, \ldots)W_0$$
$$\text{where head}_i = \text{Attention}\left(QW_i^Q, KW_i^k, VW_i^V\right) \tag{5}$$

4.4 Model

In model designing, we have followed similarities with ViT, so that this ViT 3D can be used for any problem without much alteration of the code.

Figure 4 displays the model's whole overview. The standard transformer only handles token embeddings in a 1D sequence. To properly manage 3D images, we convert the image $x_p \in R^{H*W*C*D}$ into a specific order of 2D patches $x_p \in R^{N*(p3.c)}$, , where (H, W, D) is the resolution of the original image, C is the number of channels, (P, P) is the resolution of each image patch, and $N = HWD/P^2$ is the resulting number of patches, which also serves as the effective input sequence length for the transformer. We flatten the patches and apply a trainable linear projection to translate to the transformer's constant latent vector size D across all of its layers (Eq. 1). The patch embeddings are what we call this projection's output.

The series of embedded patches ($Z_0^0 = x_{\text{class}}$), whose state at the transformer encoder's output serves as the image representation y, are prefixed with a learnable embedding that is comparable to BERT's [class] token (Eq. 4). Both during pre-training and fine-tuning, a classification head is connected to z. During pre-training, the classification head is built using an MLP with a single hidden layer, and during fine-tuning, it is built using a single linear layer.

In order to keep track of positional data, position embeddings are added to patch embeddings. Since employing more complex 2D-aware position embeddings has not resulted in any appreciable performance gains, we use conventional learnable 1D position embeddings [7]. The resulting run of embedding vectors is sent into the encoder.

Two layers in the MLP exhibit GELU nonlinearity.

$$Z_0 = \left[x_{\text{class}}; x_p^1\ \text{EM}; x_p^2\ \text{EM}; \ldots; x_p^N\ \text{EM}\right] + \text{EM}_{\text{pos}} \tag{6}$$

$$\text{EM} \in R^{(P3 \cdot C) \times D},\ \text{EM}_{\text{pos}} \in R(N+1) \times D \tag{7}$$

$$Z_{\tilde{l}} = \text{MSA}(\text{LNM}(z_{l-1})) + z_{l-1} \quad l = 1 \ldots L \tag{8}$$

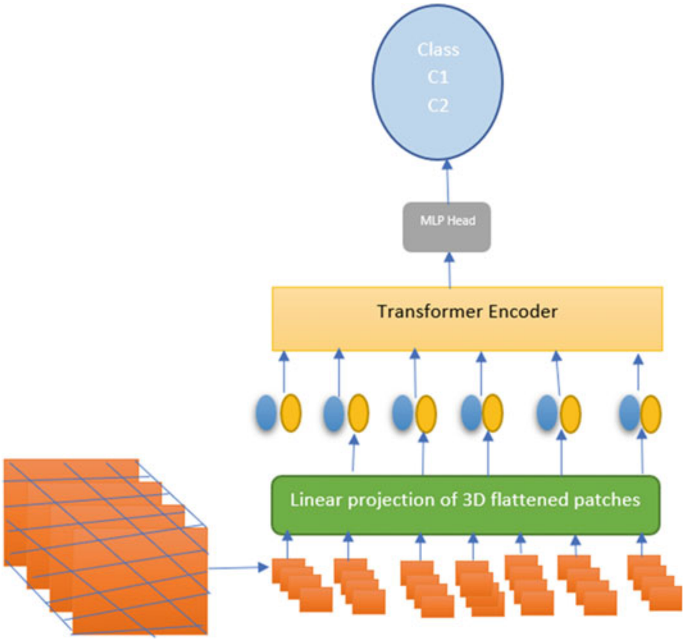


Fig. 4 ViT 3D architecture

$$Z_l = \mathrm{MLP}(\mathrm{LNM}(z_l)) + Z_{\tilde{l}} \quad l = 1 \ldots L \tag{9}$$

$$Y = \mathrm{LNM}\left(Z_L^0\right) \tag{10}$$

Our contribution enables ViT to apply in medical imaging and video analytics.

5 Performance Results

We have trained the model on the Nvidia RTX 3070 (8 GB GPU), on a fold basis and trained for 10 epochs performance at each epoch which is shown in the below table. Metrics: weighted multi-label logarithmic loss and a little about this is explained below:

You must supply a set of anticipated probabilities for each exam ID (a separate row for each cervical level subtype). Next, we calculate the log loss between each projected probability and its actual label.

For label j on exam I, the binary-weighted log loss function is defined as:

$$L_{ij} = -w_j * \left[y_{ij} * \log\left(p_{ij}\right) + \left(1 - y_{ij}\right) * \log\left(1 - p_{ij}\right)\right] \tag{11}$$

Table 1 Training loss and validation loss for different epochs

Epoch	Training loss	Validation loss
1	0.7406925	0.7339912
2	0.7304091	0.6990909
3	0.7185495	0.6987427
4	0.7200148	0.6995985
5	0.7201688	0.6979118
6	0.7193345	0.6978509
7	0.7196428	0.6977925

Table 2 DenseNet results for different epochs

Epoch	Training loss	Validation loss
1	0.764932	0.765301
2	0.765678	0.765674
3	0.765702	0.765555
4	0.766168	0.765711
5	0.766252	0.765471
6	0.751578	0.754715
7	0.751450	0.754715

Finally, loss is averaged across all rows.

The training and validation loss of the ViT 3D is given in Table 1.

We can see that our evaluation metric is partially grabbed down from 0.73 to 0.69, and this experiment setup has taken around approx. of 10–12 h of training and validation. For a comparative study of our model, we also experimented with DenseNet under a 2D environment, and the results are given below (Table 2).

As we can see here the model loss became stagnant after the fifth epoch when comparing the validation loss of the ViT 3D model and DenseNet 101 model, the difference between the loss values is so significant in such a way that our 3D model is giving more promising results when compared to DenseNet.

The more dominant transformer architecture may help us to achieve the results of ViT 3D, and there may be many models.

6 Conclusions and Further Discussions

As we can see there, few improvements can be made in this proposed methodology.

1. We can take the frames which we have stacked together as a separate entity and take their path size of them separately.
2. Using the self-attention mechanism with small variations may yield us better results.

3. Using the multi-GPU environment better than RTX 3070 for faster evaluations and betterment of the architecture or the data augmentation methodologies.
4. Experimenting with the SOTA models of image classification and performance evaluation with the proposed methodology.

References

1. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
2. Salehinejad H et al (2021) Deep sequential learning for cervical spine fracture detection on computed tomography imaging. In: 2021 IEEE 18th international symposium on biomedical imaging (ISBI), pp 1911–1914. https://doi.org/10.1109/ISBI48211.2021.9434126
3. Phal PM, Riccelli LP, Wang P, Nesbit GM, Anderson JC (2008) Fracture detection in the cervical spine with multidetector CT: 1-mm versus 3-mm axial images. Fracture detection in the cervical spine with multidetector CT: 1-mm versus 3-mm axial images. Am J Neuroradiol (ajnr.org)
4. Engstler P, Keicher M, Schinz D, Mach K, Gersing AS, Foreman SC et al (2022) Interpretable vertebral fracture diagnosis. arXiv preprint arXiv:2203.16273
5. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN et al (2017) Attention is all you need. Adv Neural Inf Process Syst 30
6. Fukui H, Hirakawa T, Yamashita T, Fujiyoshi H (2019) Attention branch network: learning of attention mechanism for visual explanation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10705–10714
7. Liu G, Guo J (2019) Bidirectional LSTM with attention mechanism and convolutional layer for text classification. Neurocomputing 337:325–338
8. Choi E, Bahadori MT, Sun J, Kulas J, Schuetz A, Stewart W (2016) Retain: an interpretable predictive model for healthcare using reverse time attention mechanism. Adv Neural Inf Process Syst 29
9. Yan C et al (2020) STAT: spatial-temporal attention mechanism for video captioning. IEEE Trans Multim 22(1):229–241. https://doi.org/10.1109/TMM.2019.2924576
10. McClenny L, Braga-Neto U (2020) Self-adaptive physics-informed neural networks using a soft attention mechanism. arXiv preprint arXiv:2009.04544
11. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T et al (2020) An image is worth 16×16 words: transformers for image recognition at scale. arXiv preprint arXiv:2010.11929
12. Cheng Z, Wan H, Shen X, Wu Z (2021) Patchformer: a versatile 3d transformer based on patch attention. arXiv preprint arXiv:2111.00207
13. Wang Y, Fan Z, Chen T, Fan H, Wang Z (2022) Can we solve 3D vision tasks starting from a 2D vision transformer? arXiv preprint arXiv:2209.07026

Crop Disease Detection Using NLP and Deep Learning

Talla Prashanthi, Tumma Susmitha, and Rupesh Kumar Mishra

Abstract Nowadays, we have seen in our state, country, and others places that lot of crops are affecting from different diseases from different insects. So we have planned to develop a system for ranking of those regions which is basically infected with several diseases and which places infected more in comparisons with other regions. This work presents the first step toward a fully built, semantically enhanced decision support system for IPM. The ultimate objective is to construct a method to aid farmers in making decisions regarding the prevention of illnesses and pests as well as to create a comprehensive agriculture expertise compiled by collecting data from several different sources. With the help of NLP, sentiment analysis of the data is known, and when the farmers give text data of the symptoms as input, the disease name is generated using approaches for machine learning.

Keywords NLP · BERT · NER · ICT · SVM · LSTM

1 Introduction

In the countless years past, the earth's global temperature has increased by 0.8 °C and many scientists said that the temperature is anticipated to arrive 6% increment by the end of the century. Nowadays, we have seen every state and other places of our country are being affected with different diseases from different insects. India, being a tropical country, is most likely to be affected with large impacts of emerging climate change.

T. Prashanthi (✉) · T. Susmitha · R. K. Mishra
CBIT(A), Hyderabad, Telangana, India
e-mail: Prashanthit_cse@cbit.ac.in

T. Susmitha
e-mail: susmithat_cse@cbit.ac.in

R. K. Mishra
e-mail: rupeshmishra_cse@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_21

India is known being the primary grower of numerous fruits worldwide, vegetables, and major spices like chili, ginger, and pepper. Agriculture is one of the most important occupations practiced in India. And the second-largest producer of wheat is India, rice and dry fruits, and many other raw resources for textiles [1, 2]. India is ranked as one of the top-five producers of meat from cattle and poultry. The Government of India has set a target to increase the agricultural production by 4%, from 2021–2022. There are also many agricultural initiatives schemes that are launched by the Prime Minister of India "Narendra Modi" like "Atal Bhujal Yojana (ABY)," "Gramin Bhandaran Yojana (GBY)," "NMSA," "Micro Irrigation Fund (IMF)," and many other for the development and improvement of agricultural activities in India. In India, one-fourth of the total population are farmers. For most people, agriculture is their primary source of income people in India. It is said that by 2050, production of crops worldwide will be increased at least 50% in order to meet the demand. Africa and Asia provide the majority of the world's food.

In India, crop diseases vary in different regions across the country due to humidity, temperature, and rainfall. Crop diseases and pests also cause the crops to be destroyed and leading to the decrement in the food security and food production. Both the small-scale and large-scale farm owners ought to be given the basic or pertinent details about the best agricultural methods to fight against the crop diseases and pests. One of the main causes of the decline in food production is the presence of pathogens, inadequate disease control, and climate change.

Many new technologies have been developed to reduce the diseases and pests, and many laboratory approaches are done for the crop disease detections, but these methods take a lot of time and are not cost-effective. In classifying the crop diseases, the traditional methods are not feasible. Modern computer vision models provide a quick, workable, and precise answer to the issues. The end goal is to create a system that will assist farmers in making decisions on the control of pests and crop diseases as well as to create a thorough agricultural knowledge base by assembling information from several heterogeneous sources.

2 Dataset Description

We manually built 2000 rows of dataset which consists of different crop types in particular, rice, corn, chickpeas, kidney beans, pigeon peas, moth beans, mung beans, black gram, lentil, pomegranate, banana, mango, grapes, watermelon, apple, and orange, and the second column consist of disease name, namely anthracnose, Canker, bacterial streaming, crown gall, and psorosis, and third column consists of all the remedies of the particular disease, and the final column consists of text data regarding the symptoms and the disease [3]. The ultimate objective is to construct a technique to assist farmers in choosing how to reduce illnesses and problems as well as to create a comprehensive agriculture expertise compiled by collecting info from several different sources. With the help of NLP, sentiment analysis of the data is known, and when the farmers give text data of the symptoms as input, the disease name is generated with the help of machine learning techniques (Fig. 1).

3 Methodology

The proposed approach mainly focuses on detecting the crop disease with the help of text data and ranking of crop diseases that are generated, with the help of NLP techniques. Here, we used BERT. Bidirectional encoder representations from transformers (BERT) is an open-source machine learning framework for natural language processing (NLP) and language understanding [4]. It is designed to help understand the computers in the different languages or text. The BERT must be pre-trained in order to understand how the language works and implements it. The pre-training can be done from Wikipedia by utilizing text from it and can be adjusted with both inquiries and responses.

3.1 Pests and Diseases Recognition

Automatic pest and disease recognition is one of the many ways that ICTs are used in agriculture. The most popular method is image processing with advanced artificial intelligence methods like deep learning. In some circumstances, information received from sensors or other inputs is added to picture processing.

The present widespread use of smartphones with high-resolution cameras is advantageous for image-based pest recognition algorithms. However, these methods only address pests and diseases that have an obvious effect plants' structural characteristics elements (flowers, fruit, stalks, and leaves). These instruments will not be able to detect other injury-related disorders that cannot be seen in a photograph (e.g., premature fruit drop). The inherent ambiguity and imprecision of natural language limits the accuracy of language-guided techniques, which are more versatile in terms of coverage. In various contexts, various semantic technologies, such as ontologies, shown to be helpful in reducing effects of ambiguity in language. Knowledge technologies are used extensively in agriculture [5]. The agronomy domain now has a number of ontologies and structured vocabularies accessible. The majority of the ontologies and vocabularies created the hosting of agronomic data representation and annotation at Agro Portal, which has emerged as the reference repository. More particularly, the authors of reference offer a thorough analysis of the application of graphs of knowledge in the field of agricultural diseases and pests (Fig. 2).

Crop Type	Disease Name	Remedies	Text
Rice	anthracnose	Create a well balanced soil	Anthracnose is a common fungal disease of shade trees that results in leaf spots,cupping or curling of leaves and early leaf drop.
Rice	anthracnose	Create a well balanced soil	Anthracnose is caused by several different, but closely related fungi.Most fungi that cause anthracnose can infect only one type of tree.
Rice	anthracnose	Understand the mechanism of infection	Anthracnose can occasionally occur on any tree in the summer. If cool, wet weather occurs when the tree is producing a new flush o0f young leaves.
Rice	anthracnose	Keep plants healthy: Proper watering, mulching, pruning and fertilizing	Anthracnose causes the wilting,withering and dying of tissues.It commonly infects the developing shoots and leaves.
Rice	anthracnose	Understand the mechanism of infection	Anthracnose, a group of fungal diseases that affect a variety of plants in warm, humid areas,Shade trees such as sycamore, ash, oak and maple.
Rice	anthracnose	Understand the mechanism of infection	Anthracnose disease is induced by the fungus Colletotrichum lagenarium and characteristic symptoms include small, yellowish watery spots
Rice	anthracnose	Keep a clean garden: roguing, rotating crops and sanitizing tools	Symptoms of anthracnose vary from host to host, but in general,include irregular spots,and dead areas on leaves that often follow the veins of the leaves.
Rice	anthracnose	Understand the mechanism of infection	Anthracnose,a group of fungal diseases that affect a variety of plants in warm humid areas.
Rice	anthracnose	Understand the mechanism of infection	Anthracnose is a common spring disease on maple trees common in the landscape including red(Acer rubrum), silver(Acer Saccharinum), sugar.
Rice	anthracnose	Keep a clean garden: roguing, rotating crops and sanitizing tools	Anthracnose is a common and widespread rot of ripe or overripe tomato fruit. Symptoms are rare on green fruit. Symptoms on ripe fruit are small, sunken.
Rice	anthracnose	Keep plants healthy: Proper watering, mulching, pruning and fertilizing	Anthracnose disease is included by the fungus colletotrichum lagenarium, and the characteristic symptoms include small, yellowish watery spots that enlarges.
Rice	anthracnose	Understand the mechanism of infection	Anthracnose is a group of fungal diseases - all fueled by excess water on leaves, stems and fruit. During dry weather, anthracnose slows.
Rice	anthracnose	Keep a clean garden: roguing, rotating crops and sanitizing tools	Anthracnose varies depending on the specific fungus and plant involved, but tender leaves are usually hardest hit.

Fig. 1 Datasets

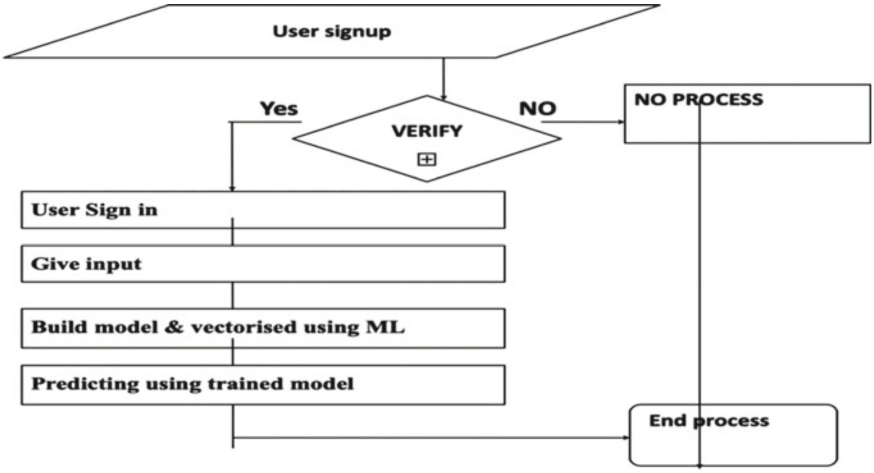


Fig. 2 Block diagram

4 Evaluation and Results Analysis

Farmers in our nation (particularly in our region) rely largely on agricultural centers for the detection and diagnosis of crop illnesses since they lack the literacy needed to learn about the crops and crop diseases on their own. Typically, photographs of crops collected from farm fields are examined in a laboratory at an agricultural research facility to identify crop diseases, after which the farmer is informed. It appears to take a lot of time. The alternative way of identifying and diagnosing plant disease through visual examination of plant leaf symptoms entails a notably high level of complexities. Even an experienced agronomist and plant pathologist frequently fails to properly recognize these troubles, the extremely vast number of crops, and the psychopathological issues with these crops. An automated approach to identify crop illnesses and their severity would be a big step forward in this area. This method detects diseases more quickly and accurately, making it easier for agronomists to identify diseases quickly and educate farmers about the existence and severity of diseases. This data is presented using the natural language processing (NLP) paradigm [6]. Voting classifiers are machine learning models that take training data from a variety of models and predict an output (class) based on the class that has the best chance of becoming the output. It simply averages the outcomes of each classifier that was submitted into the voting classifier in order to forecast the output class based on the highest majority of votes. Instead of creating individual specialized models and evaluating their correctness, the idea is to develop a single model that learns from multiple models and predicts output based on their aggregate majority of voting for each output class. Plant diseases pose a serious danger to food security because they substantially lower crop yields and impair the quality of crops that are grown as food grains. Plant protection is a key aspect of the sustainable

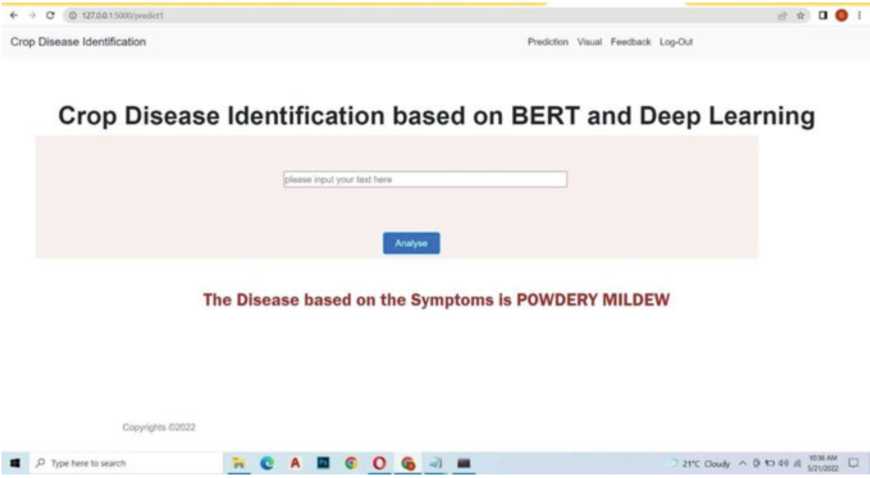


Fig. 3 Output page

agriculture challenge. Numerous studies suggest that the emergence of infections and their interactions with physiological modifications brought on by climate change. The fact that today's illnesses move throughout the world more quickly and easily than ever before makes the problem even more complicated. In areas where they were previously absent, new diseases are beginning to spread, and there is a dearth of local knowledge to combat these diseases. Without expertise, using pesticides can lead to pathogen resistance, which weakens the crops' defenses against disease. In order to prevent needless resource waste and promote healthier growth, it is crucial to address the issue of pathogen resistance. One of the major challenges in this setting is the accurate and prompt detection and diagnosis of illnesses [7]. The diagnosis of crop diseases has relied on human annotation through visual inspection. After clicking on the sign-up button, the users will be directed to sign-up/sign in page. If the user he/she is using this application for the first time, they can register/sign up in the sign-up section. Or else, the users can enter their credentials and sign in to their accounts. After clicking analyze, this page will appear, based on the symptoms descriptions, the disease name is displayed (Fig. 3).

5 Conclusion and Future Scope

Organic farming generally takes a cause-and-effect approach to problems rather than just treating their symptoms. Because of this, it is essential to identify pest or disease outbreaks as soon as possible, so that preventive actions can be taken. However, farmers typically lack the skills and resources needed to identify the precipitating events and take appropriate action. Furthermore, the majority of individuals are still unaware of therapies that adhere to organic agriculture the ability to identify the presence of pests and diseases in their crops and, in order to reduce their negative impacts, to design preventive measures and utilize IPM techniques approved for organic agriculture. In the future, ranking of crop diseases can be done using BERT architecture. Through the sentence embedding, the BERT can classify our datasets in the form of sentence only. Now after that we will apply LSTM machine leaning algorithm for convert sentence embedding into regression part. After doing this regression, we have to change the regression into some fixed score for our complete sentence into region wise datasets with the help of linear kernel. Based upon unique score, we will provide the rank of region-wise crop diseases.

References

1. Gounder S, Patil M, Rokade V, More N (2021) Agrobot: an agricultural advancement to enable smart farm services using NLP. J Emerg Technol Innov Res. Available at SSRN: https://ssrn.com/abstract=3890591
2. Li L, Zhang S, Wang B (2021) Plant disease detection and classification by deep learning—a review. IEEE Access 9:56683–56698. https://doi.org/10.1109/ACCESS.2021.3069646
3. Rodríguez-García MA, García-Sánchez F, Valencia-García R, Knowledge, based system for crop pests and diseases recognition. https://doi.org/10.3390/electronics10080905
4. Geetha S, Balajia S, Santhiya, Subicsha, Subashri, Farm's smart BOT. https://doi.org/10.17762/turcomat.v12i10.4986
5. Nagaraju M, Chawla P (2020) Systematic review of deep learning techniques in plant disease detection. Int J Syst Assur Eng Manag 11:547–560. https://doi.org/10.1007/s13198-020-00972-1
6. Vanib V, Guruprakash CD (2022) A review on smart agricultural applications: crop yield and plant disease prediction. In: 2022 International conference on applied artificial intelligence and computing (ICAAIC), pp 885–892. https://doi.org/10.1109/ICAAIC53929.2022.9792736
7. Ayub U, Moqurrab SA (2018) Predicting crop diseases using data mining approaches: classification. In: 2018 1st international conference on power, energy and smart grid (ICPESG), pp 1–60. https://doi.org/10.1109/ICPESG.2018.8384523

Design of a Virtual Mouse for Hand Gesture Recognition and Characterization

Shaik Salma Begum, K. Bala Brahmeswara, Appala Yochana, and K. Durga Prasanth

Abstract One of the marvels of human–computer interaction (HCI) technology is the mouse. Since a wireless mouse or Bluetooth mouse still need a battery for power and a dongle to connect it to the PC, they are not entirely device free at this time. This restriction can be removed by using computer vision to capture hand movements and hand tip detection using a webcam or built-in camera. Without using a real mouse, the computer can be virtually controlled using hand gestures to accomplish left click, right click, scrolling, and computer cursor tasks. As a result, the suggested method will stop the spread of COVID-19 by doing away with human interaction and the need for external devices to operate the computer. A replacement of conventional mouse system is done by this proposed system and to perform various mouse functions as well as provide control. The web camera records the hand gestures, hand tips, and later processes the frames to perform the specific functions of mouse such as drag and drop, double click, scrolling, multiple item selection, volume control, brightness, left click, and right click. To do this, we'll use PYTHON libraries like MediaPipe and comtypes, as well as OpenCV and CNN for hand aspect ratios and hand detection, respectively.

Keywords HCI · Background subtraction · Color detection · Gesture-controlled mouse · Image processing

S. S. Begum (✉) · K. B. Brahmeswara · A. Yochana · K. D. Prasanth
Department of Computer Science and Engineering, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru 521356, India
e-mail: shaiksalma.gec@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_22

1 Introduction

1.1 Introduction

Our society cannot exist without computers and other electronic devices. They have had a significant impact on us. Communication is one of the main effects of computers. Despite the enormous advances in computing, most human–computer interactions (HCI) still rely on input devices like keyboards and mouse. The devices we use every day are getting smaller thanks to the advancement of Bluetooth or wireless technologies, as well as advancements in the field of augmented reality. When utilizing a wireless or Bluetooth mouse, several accessories are needed, including the mouse, a dongle to connect to the computer, and a battery to power the mouse. The most effective and expressive form of human communication, and one that is widely understood, is hand gesture. It is sufficiently expressive for both the deaf and the dumb to understand it. Since humans first began utilizing hand gestures to communicate in society, many generations have passed. The thumbs-up and thumbs-down signals have always been present in the environment. Gestures are said to be the simplest form of communication.

In today's automated world, hand gesture technology is used in a wide range of sectors, including banking, industry, IT hubs, and medical applications. A hardware and software system called the human–machine interface (HMI) facilitates information sharing and communication between humans and machines. We frequently use a variety of indicators, including LEDs, switches, touch screens, and LCDs, as part of HMI devices. Another unusual method of communicating with machines like robots or computers is through hand gestures. Ever wanted to use your laptop or computer by merely waving your finger in the air? We discuss the difficulties of utilizing gestures to operate computer programs that have both static symbols and moving objects. Either static model data or dynamic systems with linear parameters are used to model each gesture. We may now easily use hand movements and hand gestures to operate the functions of the laptop instead of utilizing a variety of devices like keyboards, mouse, joysticks, and so on. This concept is based on the widely used belief that controlling a laptop or computer may be done using hand gestures.

In this paper, we employed hand gesture control, which enables users to manage a wide range of features, including navigation, music, video, documents, games, and more. Only a little amount of processing and memory are needed for real-time recognition. We look into the proper motions, how to spot them, and which orders they should override. Our tablets, iPads, laptops, and smart phones all have a variety of technical breakthroughs, including facial identification, biometric authentication, and natural language processing.

The mouse pointer might be moved in hand movement identification, a more recent method of human–machine interaction, by simply positioning our figure in front of the computer's web camera. This paper could be used for completing computer mouse pointer and scroll functions without the usage of a typical mouse device by using a web camera or a built-in camera in the computer [1–4].

1.2 Problem Statement

There may be instances in the real world, where a physical mouse cannot be used due to lack of space or for people with hand issues who are unable to use a physical mouse. Additionally, given the COVID-19 situation, it is not safe to use electronic devices by touching them because doing so could lead to a scenario, in which the virus is spread by contacting the gadgets. Because the suggested system uses hand gesture and hand tip detection to operate PC mouse functionalities via a webcam or built-in camera, it can be used to get around these issues. Some gadgets, including the mouse and the dongle to connect, must be used when utilizing a wireless or Bluetooth mouse.

The general issue with the physical mouse as it currently exists is as follows:

- Mechanical deterioration occurs with real mouse.
- A certain piece of hardware and surface is needed to utilize a physical mouse.
- The performance of a physical mouse varies depending on the situation and is difficult to adapt to.
- Both wired and wireless mouse have their own life expectancies.

1.3 Existing System

- **Wired Mouse**: A wired mouse delivers data via the cord and is directly connected to your desktop or laptop, typically through a USB connection. The cord connection offers a number of significant benefits.
- **Wireless Mouse**: A receiver attached to your computer receives radio signals from wireless mouse. When a signal is received, the computer interprets it to determine how the cursor was moved or which buttons were pressed. There are various limitations despite the freedom or range of wireless models.
- **Bluetooth Mouse**: A Bluetooth mouse uses your computer's inherent Bluetooth connection to connect to numerous devices simultaneously (Fig. 1).
- Some current virtual mouse systems use hand gesture detection by donning a glove for the hand and also recognize gestures by the color of the tips of the fingers (Fig. 2).

1.4 Disadvantages

- Wearing gloves makes recognition less accurate, and in some situations, failing to detect color tips also makes recognition less precise.
- RGB-D cameras are used in the virtual mouse that is now available and uses fingertip detection.

	Wired Mouse	Wireless Mouse	Bluetooth Mouse
Advantages	• Better accuracy • Low-latency • More affordable • No batteries required	• Extended range from computer • Not restrained by cord length • Unrestricted movement • Doesn't require mousepad	• Extended range from computer • Not restrained by cord length • Unrestricted movement • Doesn't require mousepad
Disadvantages	• Less freedom of movement • Constrained to length of cord • Less convenient	• Requires batteries to operate • Lower accuracy (may be negligible for non-gamer users) • More expensive than wired models	• Requires batteries to operate • Lower accuracy (may be negligible for non-gamer users) • More expensive than wired models • May require purchase of USB Bluetooth receiver if your machine does not have Bluetooth functionality

Fig. 1 Comparison between different kinds of mouses

Fig. 2 Virtual mouse implementation using color tips

- Although the proposed system by Quam produces more accurate results, it is challenging to use the system for some gesture controllers and a DataGlove is needed.
- The "cursor control via hand gestures" system is built on a system of several color bands, where various colors correspond to various functions. The key to using the mouse is the amount of colors, yet the system was controlled by a variety of colors. Instead than using different movements, a function is performed by the amount of colors.
- Although it is based on background extraction and contours identification, the "hand gesture recognition for human–computer interaction" technology is quite slow to use.

1.5 Proposed System

The suggested method uses hand gestures to simplify human–computer interaction. The amount of direct contact with the computer is minimal. Dynamic hand gestures may essentially control all input and output processes. This proposed method uses cutting-edge machine learning and computer vision algorithms to recognize hand movements, and it does so without the use of any special hardware. It makes use of models like convolutional neural networks (CNNs), in which MediaPipe, which is built on top of pybind11, has incorporated.

The primary goal of the proposed virtual mouse system is to create a replacement for the conventional mouse system to perform and control mouse functions. This can be done with the aid of a web camera that records hand gestures and hand tips and then processes these frames to perform the specific mouse function, such as the left click, right click, and scrolling function. The MediaPipe framework is utilized for hand tracking and gesture detection, and the OpenCV library is used for computer vision. To track and identify hand movements and hand tips, the algorithm uses machine learning ideas.

A package for computer vision called OpenCV has image-processing tools for object detection. Real-time computer vision applications can be created by utilizing the OpenCV library for the Python programming language. The processing of images and videos as well as analytical techniques like face and object detection relies on the OpenCV library.

A framework called MediaPipe is a Google open-source framework that is applied in a machine learning pipeline. Since the MediaPipe framework was created utilizing time series data, it can be used for cross-platform programming. Real-time detection and identification of a hand or palm are accomplished using a single-shot detector model. The MediaPipe uses the single-shot detector model. Because it is simpler to learn palms, the hand detection module initially trains a model for palm detection. Furthermore, for small objects like hands or fists, the non-maximum suppression performs noticeably better.

A convolutional neural network (ConvNet/CNN) is a deep learning method that can take in an input image, give various elements and objects in the image importance (learnable weights and biases), and be able to distinguish between them. Comparatively speaking, a ConvNet requires substantially less preprocessing than other classification techniques. ConvNets have the capacity to learn these filters and properties, whereas in primitive techniques, filters are hand engineered (Fig. 3).

The ability of artificial intelligence to close the gap between human and machine skills has dramatically increased. Both professionals and amateurs focus on many facets of the field to achieve great results. The field of computer vision is one of several such disciplines.

The goal of this field is to give computers the ability to see and understand the world similarly to humans do. They will then be able to use this understanding for a variety of tasks, including image and video recognition, image analysis and classification, media recreation, recommendation systems, natural language processing, etc. With

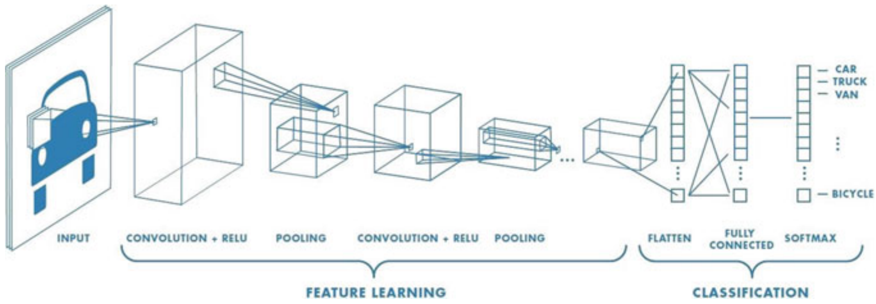


Fig. 3 Picture demonstrating the working of CNN algorithm

time and a focus on one specific algorithm—the convolutional neural network—deep learning for computer vision has advanced and been refined.

1.6 Advantages

- The suggested technique can function correctly in every lighting situation and for skin tones of any color.
- Easily navigable.
- Transportable.
- Hardware is free.
- Basic hand motions.
- Real-time software.

2 Components

Since this paper is a prototype for all computers, specific components cannot be used. As a result, the following requirements must be met:

2.1 **Equipment**: Webcam A webcam is a requirement for identifying the image. The relationship between mouse sensitivity and camera resolution is direct. An improved user experience is assured if the camera's resolution is sufficient. Every time the computer turns on, the webcam captures images in real time. The system will decide the appropriate action based on gestures and finger motion.

2.2 **Software**: Support for modules and packages in Python promote modularity in programs and code reuse. For all popular platforms, the Python interpreter and the comprehensive standard library are freely distributable and available in source or binary form.

3 Mechanisms

There are 4 modules in the software:

A. **Hand-tracking Module**: Before beginning, we turn on the webcam for video recording. In order to capture everything, the webcam records one frame each second. The system first reads the frame before changing the frame's color. Next, the width and height of the hand landmarks' coordinates are determined. After that, the coordinates are transformed into pixels. Thus, using colored counters to demonstrate the hand landmarks.

B. **Volume Control**: This module's purpose is to enable hand gestures to control the computer's volume. In order for the system to record, analyze, and interpret the user's hand gestures and determine the coordinates of the hand landmarks, we first import the hand-tracking module. The system's minimum and maximum volume ranges are then set. The landmark coordinates are then set in relation to the previously stated volume range. Therefore, the computer's loudness will depend on the distance between the user's thumb and forefinger.

C. **Virtual Painter**: With the help of this module, users can use their fingers to make drawings on the screen. Once more, we import the hand-tracking module first. To make it more interactive, we need a few header files. We made advantage of a few previously existing hand-tracking module routines. Then, we specify two modes: drawing mode and selecting mode. The user can choose the color of the brush in selection mode, and they can draw on the board in drawing mode.

D. **Mouse Control**: Using just our first two fingers, we can operate the computer mouse and do simple mouse tasks with the help of this module. We make use of the Python language's built-in library "autopy" in this module. This library gives the user the ability to use mouse actions quickly and easily. With the help of this module, the user is able to move, scroll, and click the mouse using simple hand gestures.

4 Design

4.1 **Video Capturing**: In order to interpret hand gestures, we must first record real-time video of the user's movements. This is accomplished by employing a webcam, which continuously outputs a series of photos at a specific frame rate (Frames per second) (FPS).

4.2 **Calculation of hand landmark coordinates**: After collecting the gestures, the system analyzes the images and determines the hand landmark coordinates.

4.3 **Tracking the Cursor**: The mouse driver is accessed after the coordinates have been determined, and the coordinates are then passed to the cursor. These coordinates are used to move the cursor to the proper location. As a result, the mouse follows the user's hand movements across the camera's range of vision accordingly (Fig. 4).

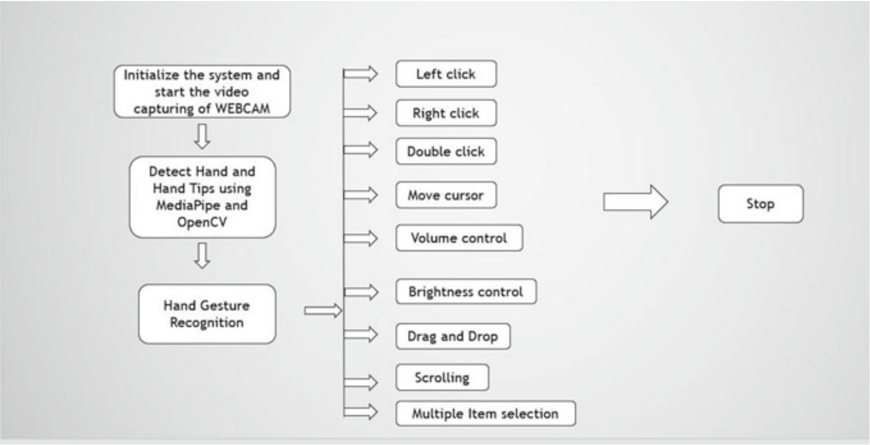


Fig. 4 Activity diagram for gesture-controlled virtual mouse

5 Implementation

(i) Open the Anaconda Prompt.
(ii) Change the location of prompt to destination file with help of commands.
(iii) Now type "Jupyter notebook" and press enter
(iv) Then Jupyter notebook will open with files at the desired location.
(v) Now click on the file need to be executed.
(vi) It will then open the file in the Jupyter notebook.
(vii) The executable file will be opened (Figs. 5, 6, 7, 8, 9 and 10).

Fig. 5 Executed output

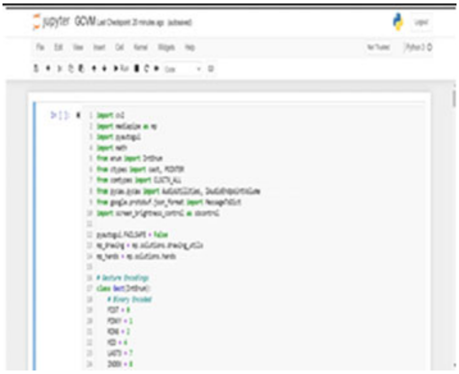


Fig. 6 "GCVM.ipynb" in Jupyter notebook

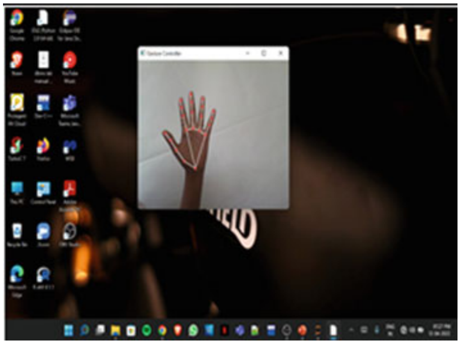


Fig. 7 Neutral gesture: used to halt/stop execution of current gesture

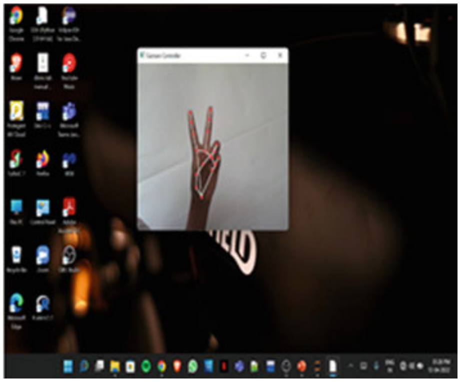


Fig. 8 Move cursor: cursor is assigned to the midpoint of index and middle fingers. This gesture moves the cursor to the desired location

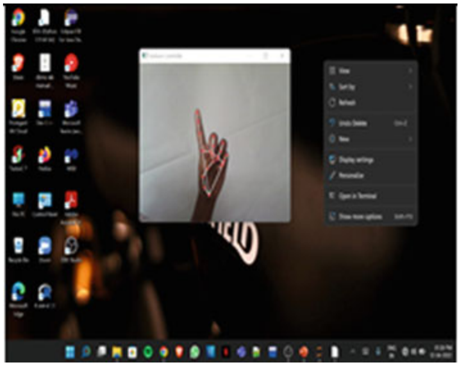


Fig. 9 Right click: gesture for single right click

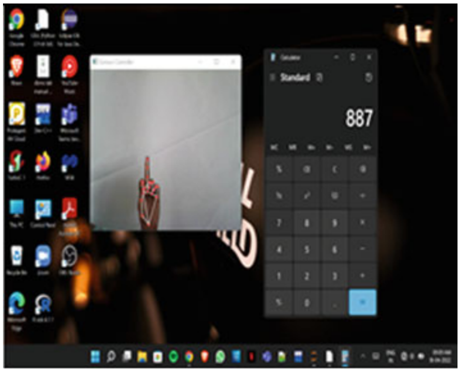


Fig. 10 Left click: gesture for single left click

6 Results and Discussions

6.1 **Hand-Tracking Module:** We can see that the video is being recorded and is being displayed in an another window. The hand above the wrist is shown in the video, and it has 21 points with a line connecting them all. Different colors are used to highlight the points and the line connecting them, respectively (Fig. 11).

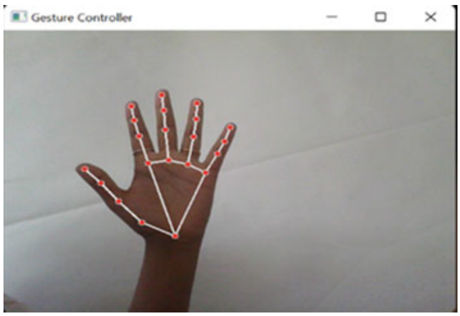


Fig. 11 Hand-tracking using the algorithm

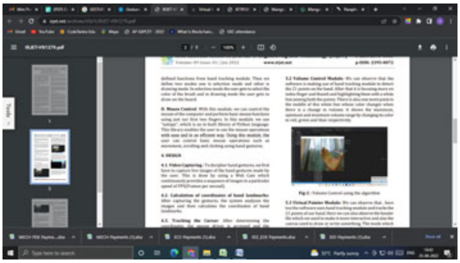


Fig. 12 Volume control using the algorithm

6.2 **Volume Control Module**: We can see that the software uses a hand-tracking module to find the hand's 21 locations. The index finger and thumb are then given extra attention, and a white line connecting the two spots highlights them. Additionally, there is another point in the center of this white line, and its color varies in response to variations in volume. It changes to red, green, or blue to indicate the maximum, ideal, and minimum volume ranges (Fig. 12).

6.3 **Virtual Painter Module**: We can see that the software is tracking our hand's 21 points in this instance as well using a hand-tracking module. We can see the header file that we used to make it more interactive and the canvas that we can use to write or draw on. The mode that we are using, such as the drawing mode or the selection mode, is immediately printed on the terminal (Fig. 13).

6.4 **Mouse Control Module**: As we move our index finger on the screen, the pointer on the screen moves as well. This module's operation and display window are very similar to the hand-tracking module. This demonstrates how the cursor can be moved with the index finger (Fig. 14).

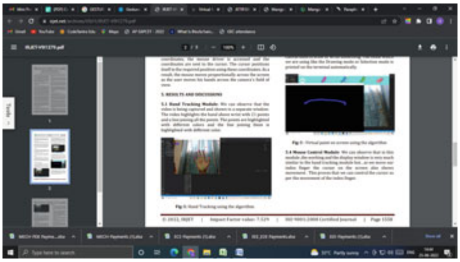


Fig. 13 Virtual paint on screen using the algorithm

Fig. 14 Mouse cursor control using the algorithm

7 Conclusions and Future Scope

7.1 Conclusion

The most effective human–machine interface is provided through gesture recognition. The development of alternative human computer interaction modes depends on gesture recognition. It makes it easier for people to interact more naturally with machines.

7.2 Future Scope

A wide range of industries, including augmented reality, computer graphics, gaming, prosthetics, and biomedical instruments, can benefit from this technology. In a variation of our system known as Digital Canvas, which is gaining popularity among painters, the artist can use the hand as a brush and the virtual mouse technology to produce 2D or 3D images, with a virtual reality kit or a monitor serving as the display set. Patients who lack control over their limbs can benefit from this device. Modern gaming consoles have incorporated computer visuals and gaming technology to create interactive games that track player actions and translate them into commands.

References

1. https://www.hindawi.com/journals/jhe/2021/8133076/
2. https://www.researchgate.net/publication/336304431_Design_and_Development_of_Hand_Gesture_Based_Virtual_Mouse

3. Begum SS, Rajya Lakshmi D (2019) GLCM of fuzzy clustering means for textural future extraction of brain tumor in probabilistic neural networks. Int J Innov Technol Explor Eng 9(1). ISSN: 2278-3075
4. Kadaru BB et al (2018) A novel ensemble decision tree classifier using hybrid feature selection measures for Parkinson's disease prediction. Int J Data Sci (IJDS) 3(4). ISSN:2053-082X

Voice-Activated System for Disabled People Using IoT

Farah Anjum and M. Seetha

Abstract The Internet of Things (IoT) is a growing and disruptive technology with standards emerging for wireless communication between sensors, actuators, and gadgets in everyday human life. It is manifest that the medical sector needs a technology system for providing health care to persons with impairments amid the rise of IoT systems. This paper describes a voice-activated bed for bedridden individuals. The proposed method includes an EasyVR Shield 3 voice recognition module, an Arduino UNO, a servo motor, and an adjustable bed. Voice recognition modules must be programmed before they comprehend commands. After correctly identifying the voice command, Arduino sends the flow to the servo motor. Adjustable beds can be elevated or lowered in four different ways, depending on the user's comfort and needs. It is also possible to convert the bed into a chair via voice commands. This study aims to make bedridden patients more comfortable and confident to live independently at home, minimizing the need for additional help.

Keywords Smart bed · Arduino UNO · Voice recognition · Sensors · Adjustable bed

1 Introduction

Health monitoring is essential as such our daily activities. For elderly, blind, disabled, and patients living in their own homes without assistance, numerous new solutions have been developed in the recent years to allow them a means to be independent. The main issue for bedridden is the need for localization or assistance from others. Nowadays, hospitals are initiating the integration of a variety of specialized sensors to enhance patient outcomes and overall construction efficiency. As more and more items are on the Internet, medical industries must evolve. Doctors can remotely

F. Anjum (✉) · M. Seetha
Department of CSE, G. Narayanamma Institute of Engineering and Technology, Hyderabad, Telangana, India
e-mail: farahanjum9095@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_23

monitor patients and administer medications based on monitored data with the aid of Internet tools. The difficulties of preserving mobility and cognitive performance pose challenges for older persons with special needs and spinal injuries to continue living alone and gaining some degree of autonomy. The hospital bed is for people who need hospital care while undergoing medical treatment. Over time, the hospital bed has evolved from a straightforward bed made of raw stretchers into a bed with various features and several functions [5]. The control of the bed can be made much easier by using speech recognition to operate IoT-connected devices [5]. The bed is retrofitted with sensors and motors controlled via voice commands. A microphone transmits speech commands to the control bed. The emergence of a speech recognition system had to tackle issues of distance, safety, and noise.

2 Literature Survey

This research looks at a medical care bed that uses Internet of Things technologies. It intends for hospital patients or other people who require treatment which is controlled by a button, voice commands, or phone apps. Adjustable height for the entire bed, head and feet, temperature, pressure, voice command, and programs to run both families using sensors and monitoring the patient's body temperature is just a few of the features. This set of characteristics is unique in that it addresses patient convenience and comfort while also addressing medical professionals' comfort [1].

A smart bed is a medical bed that is part of a more secure patient-care environment. A motor driver circuit connects the Raspberry Pi to a stepper motor, which is then brought forward with an audio input. The accuracy of detecting voice commands was discovered to be problematic. Wireless communication can be inconvenient for patients [2].

About home automation systems, there have been numerous studies and innovations. The voice recognition-based home automation system recognizes uttered inputs using a PC running Microsoft Speech API. These instructions are transmitted to the controller by the RF transceiver, which then controls the different electrical equipment. The addition of a computer increases the system's price and complexity [3].

A voice recognition module called the SR-07, an Arduino controller, a wheelchair, and a navigation module is used in the intelligent home navigation system for elderly and disabled people. The voice recognition module sends a command to the Arduino, which then moves the wheelchair in accordance. It eliminates the requirement for third-party help [4].

3 Proposed System

In this paper, a model is proposed for bedridden patients. Mobility of patient can be done using voice commands. A voice recognition system is built for disabled patients that provide a sense of comfort and an easier life for patients and caregivers as well. The proposed model consists of a speech recognition module, an Arduino, servo motors, and an adjustable bed. Using voice commands, patient can move the bed in four different height positions. Bed can also be converted into chair position according to their convenience. It can also be used for the patients who suffer from heart stroke and respiratory disease, in which the patient must keep their head position high. The patient who has undergone any surgeries and cannot move by themselves can move their bed position into chair without any assistance.

4 Objectives

- To train the module for voice recognition.
- To provide the movement of bed through voice commands.
- Analysis of the accuracy of the speech recognition module in four bed height modes.

5 Methodology

This system provides healthcare services for patients, the elderly, and others with special needs. The movement of the bed is operated via voice commands using IoT technology and a speech recognition module. This relies mainly on IoT. The sensors and components are embedded into the bed using IoT technology. The following components comprise the voice-enabled bed:

(a) EasyVR Shield 3.0
The EasyVR 3.0 Shield is an add-on shield that enables you to voice commands to control an Arduino. It is a multi-use speech recognition module produced to give virtually any application the ability to recognize speech in distinct situations. It is reliable, impactful, and reasonably priced. It is the third generation of the popular VRbot module and expands on its predecessor's features and functionality. With the shield's additional audio line-out/headphone connection, EasyVR 3.0 features include up to 32 user-defined speaker-dependent (SD) commands and 26 pre-built speaker-independent (SI) instructions that are ready to use. It also has access to the EasyVR module's I/O pins. Figure 1 depicts the representation of EasyVR Shield 3.0.

Fig. 1 EasyVR shield 3.0

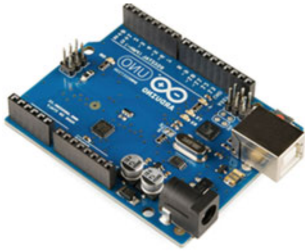


Fig. 2 Arduino UNO

(b) Arduino UNO
A microcontroller board Arduino UNO is premised on the ATmega328P. It contains six analog inputs, a 16 MHz ceramic resonator (CSTCE16M0V53-R0), a USB port, a power jack, ICSP header, and a reset button. It also has 14 digital input/output pins, where six are PWM outputs. Figure 2 shows an illustration of an Arduino UNO.

(c) Servo Motor
A servo motor is a kind of motor that has high precision rotational capabilities. It primarily consists of a control circuit that provides insight into the motor shaft's current position and enables accurate rotation. It is just a standard motor that operates via a servo mechanism. A servo motor often has a gear configuration that allows us to produce a very high torque servo motor in small and light designs. These characteristics have led to its employment in various applications, including toy cars, RC helicopters, planes, robotics, etc. Figure 3 shows a representation of a servo motor.

Fig. 3 Servo motor

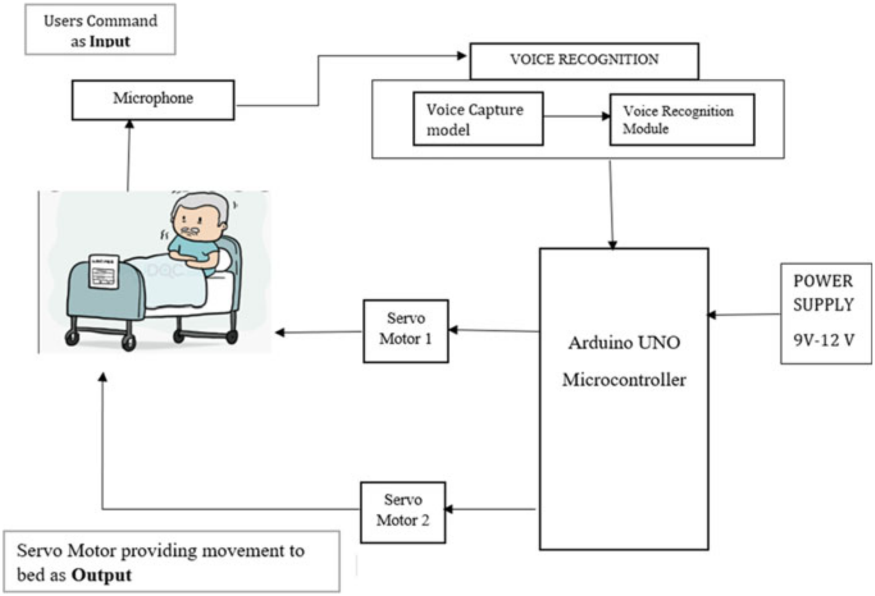


Fig. 4 Schematic diagram for the movement of bed

6 Implementation

The vision is to create a bed for homes, hospitals, and medical centers and be trusted to perform spoken commands. There are four commands to move the bed: down, rest, wake and sit. Initially, voice module must be trained with the voice commands. The microphone picks up the user's voice orders and converts them into electrical impulses. The voice module receives electrical impulses and digitizes and stores them as templates to represent user commands. When a user gives command to the system, the speech module checks whether it matches with the one that has previously been saved and sends the output to the microcontroller if it does. Then servo motor will perform the specified function. Figure 4 represents the schematic diagram for the proposed system.

7 Results

The prototype of the voice automated bed is shown in Fig. 5. This bed has fundamentally reduced the time it takes for a caretaker to help a patient. Constant observation is a massive and crucial aspect of a patient's health. It may mean that observing and being present with a patient can be responsible for enhancing or spoiling confidence in being independent. The interactions between the user's vocal instructions,

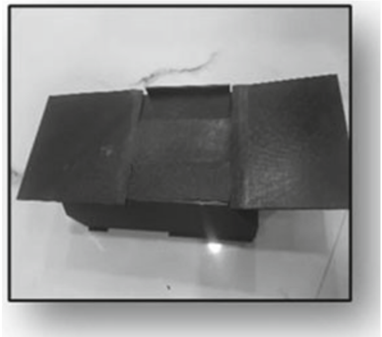


Fig. 5 Prototype of the developed bed

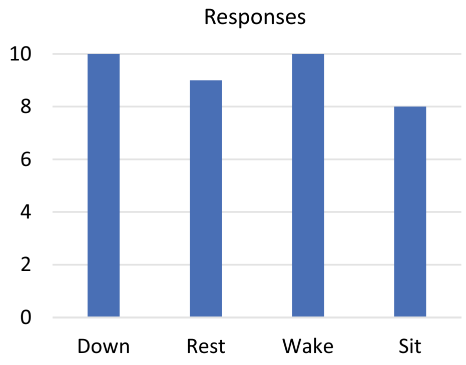
Table 1 Interaction of user voice commands and servo motors for the movement of the bed

Commands	Servo motor 1	Servo motor 2
Down	0	0
Rest	20	0
Wake	30	60
Sit	45	90

Arduino, and motors are embedded in the bed to assess the performance. Table 1 gives the outcomes of the interaction between the microcontroller and motors upon activation by the user's voice commands.

The system begins to operate only when the designed system is supplied with voltage. The machine then enters standby mode, while it waits for a voice instruction to activate it. Both servo motors (1 & 2) will be at 0° during the down command. For the rest command, servo motor one will be at a 20° angle. Servo motors 1 and 2 will be at 30° and 60°, respectively, for the wake command. Servo motors 1 and 2 will be at 45 and 90° for the sit command. Voice commands' average response time is computed. Following the test and the arithmetic mean, the achievable success of the voice commands, which reflect 85% correctness of the total orders, was determined. Figure 6 depicts the average response of the commands from user.

Fig. 6 Performance of bed



8 Conclusion

This bed will meet global demand, benefiting all patients and the elderly. Many medical situations necessitate the use of hospital beds, such as placing the body to relieve pain, avoid respiratory infections, sores, and cramps, or the need to raise the bed head due to congestive heart failure, suction, or chronic lung disease. Hospitals are using new resources in patient care. The family is intended to adjust the patient's movement, lowering the risk of bed ulcers, and lung infections. It takes some of the heavy lifting traditionally done by nurses, ultimately saving lives, and providing greater independence for older people and the disabled. The automatic bed can make life a bit easier for nurses, with 85% suffering from back injuries by having to deal with patients. Similarly, the efficiency of resource utilization, nurses and medical staff with decreased human efforts, lower costs, and bring productivity.

The future enhancement of this project is the bed can be converted into a wheelchair and operated through voice commands. Through voice commands, the patient can turn either right or left side. Patients will be more comfortable if a commode system is integrated into the bed. An advance future bed can help caregivers as well as patients more ease in their daily activities.

References

1. 19th International conference on sciences and techniques of automatic control & computer engineering (STA), Sousse, Tunisia, 24–26 March 2019
2. Ghersi I, Marino M, Miralles MT (2018) Smart medical beds in patient-care environments. BMC Med Inform Decis Mak 18:63 (IEEE)
3. Kumar M, Shimi SL (2015) Voice recognition based home automation system for paralyzed people, vol 4, no 10, pp 2508–2515
4. Kirankumar T, Bhavani B (2013) A sustainable automated system for elderly people using voice recognition and touch screen technology, vol 2, no 8, pp 265–267

5. Sakthivel K, Thavamani P, Chandrasekar R, sheik mydeen Ismail M, Govindaraj V, Sankaran S (2020) International conference on communication and signal processing, India, July 28–30, 2020
6. Voice actuated hospital bed. Int J Recent Technol Eng 8(4). https://doi.org/10.35940/ijrte.C5882.118419
7. Rajput S, Ghasti S, Giramkar V, Jadhav A, Android based Smart Bed. https://doi.org/10.5281/zenodo.3525696

SLRNetwork: Utilizing Skip Connections for Better Feature Extraction in Recognition of Sign Languages to Help Mutism

Srijarko Roy, Ankit Mathur, and Velliangiri Sarveshwaran

Abstract Through the use of sign language, persons with special needs who have oral and aural impairments have discovered a way to communicate and express their thoughts to society. However, simply knowing that sign language originated is insufficient for good dialogue and communication. We introduce SLRNetworks, a novel implementation of SLR (Sign Language Recognition) that makes use of CNNs customized along with improved feature map extraction by using the skip connections present in residual networks (ResNet). SLRNetworks will help efficiently classify the "signs" and predict their meanings, hence enabling a larger segment of society to use sign language. ResNet-18 network is used in the project's implementation as the feature extractor, while FCNs with Swiss (Softmax) activation function are used for multiclass classification. The sign language gesture images data set has been used for training and validation using the Tensorflow framework in order to ensure highly accurate results.

Keywords Computer vision · Vanishing gradient · Feature extraction · Image classification

1 Introduction

Communicating with someone who has a hearing impairment makes speaking with them difficult since they are unable to hear their counterpart and cannot communicate auditory information via speech alone. For persons who have special needs, sign language is a way to mechanically communicate in addition to speaking and hearing. People who are unable to communicate vocally or audibly have long utilized

S. Roy (✉) · A. Mathur · V. Sarveshwaran
SRM Institute of Science and Technology, Kattankulathur, India
e-mail: sr8962@srmist.edu.in

A. Mathur
e-mail: am9964@srmist.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_24

Fig. 1 Representation of various signs with their meanings

sign language to do so. CNNs for extracting feature maps and dense for image classification are used in DL and CV for SLR. A significant societal advantage would result from the use of deep learning and computer vision for sign language recognition because it would enable persons with special needs who are oral or auditory to express their thoughts and feelings to the rest of the world. Through its implementation, the needs and views of people with special needs are acknowledged and the gap between them and the general population is closed (Fig. 1).

2 Related Work

Numerous sign language classification implementations have been the subject of thorough study utilizing a wide range of data sets and approaches, each with advantages and disadvantages of its own. A 6-layered proprietary convolutional neural networks architecture trained on a custom data set under diverse circumstances, like illumination and distance, was employed in deep-ASLR [1]. It achieved an accuracy score of 99.38% using its unique data set and bespoke CNN architecture. In earlier techniques, like [2], CNNs with Hidden Markov Models (HMMs) fed with hand pictures clipped for motions were combined with an EM-based algorithm. The same authors who conducted the research in [2] expanded on it in [3], substituting glosses for cropped hand motions as the goal values. In order to improve performance, the research in [3] was further extended by integrating an LSTM atop the CNN-Hidden Markov Model architecture. For CSLR and CNN with temporal convolutions, RNNs and LSTMs have also been employed in [4], reaching 38.7% WER. Alignment networks are used in some further enhancements, as seen in [5], which aids in obtaining a WER score of 36.5%. Multi-modal learning [6] uses marginal advancements over [5] and achieves around 96.5% accuracy. SLR with temporal segmentation [7] is another method that has been successful in achieving about 82.7% accuracy. Pigou et al. employed the Italian hand gestures data set extensively to apply convolutional neural networks, and they also introduced the use of the intersection over union as a metric to assess the efficacy of their network. For the categorization of sign language, Yang Quan recommended using support vector machines (SVM) rather than intricate deep learning architectures in his work [8], which had a 99.49% accuracy rate on the Chinese manual alphabet data set. While using a Brute Force method, [9]'s usage of BoostMap Embeddings enhanced training speed by 800 times despite all models up to that point having a high time complexity. The goal of Athitsos V. et al. was to offer a quicker method for sign language recognition rather than to strive to increase measures like accuracy or word error rate per se. Rana et al. first presented unsupervised learning techniques like principal component analysis in the study [10] on extended Yale B. However, the principal component analysis yielded an accuracy only of 67.48%. On the PETS data set, graph parsing was first introduced in [11], especially ETPL(k) graph parsing, which had an accuracy of 85.4%. Continuing the application of PCA, Kamari A. et al. in their study [12] implemented MPCA-LV on the Weizmann face database for sign language recognition, marginally increasing the efficiency, and attaining an overall accuracy of 69.64%. All DL implementations of SLR have the issue of vanishing gradients. In this paper, we have implemented SLRNet which overcomes the problem, thus higher accuracy results are obtained.

3 Methodology

SLRNetwork, as illustrated in Fig. 2, uses the Residual Blocks as the feature extractor and a single FCN with a dropout of 0.6 as the final block for image classification. With a (7 × 7) kernel and a stride of value 2, the initial block (i) convolutionally processes input data. Batch normalization, activation function—ReLU and a spatial max-pooling operation with a (3 × 3) filter and a stride of value 2 are performed after the convolution operation. Residual Blocks for the primary feature extraction come after the (i). Figure 2a depicts them as Encoder Blocks (i) and Fig. 2b depicts the Residual Block's individual layers in great precision. A (3 × 3) filter is featured along with convolution with stride in each Encoder Block along with skip connections. They are in task of resolving the Vanishing Gradients problem, which impairs the performance of earlier suggested models. The network includes alternative shortcut pathways with the use of skip connections so that the cost measured at the final layer might find a different route in the event that they aren't backpropagated to earlier levels. We suggest using these skip connections through SLRNetwork to enhance the feature extraction procedure and then train the model for a possibly greater accuracy. The ResNet-18 feature extractor is used to extract the features from the SLRNet encoder, which are then given to the flattening layer to be converted into a long vector and then passed into the fully connected dense network. The primary classification duty is carried out by the fully connected network. A dropout and a single fully connected dense layer make up the FCN.

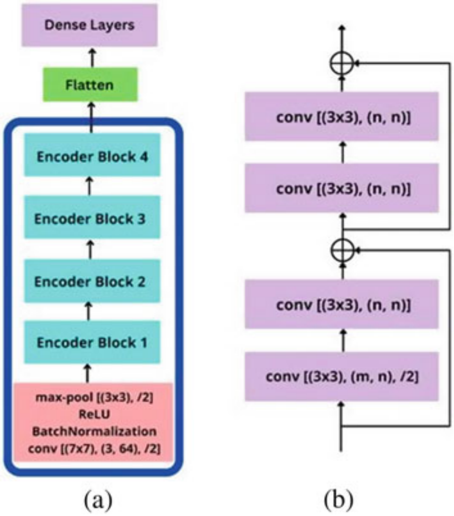


Fig. 2 SLR network architecture

To prevent overfitting, a dropout layer of 0.6 is used before the features are sent to the FCN. For multiclass classification, the dense layer comprises 37 nodes with 'Swiss' activation functions.

4 Result Analysis

4.1 Data Set

The images of sign language expressions, 55,000 jpg image data altogether, divided into 37 distinct classes, were utilized as the data set for training and validation. 70% of the data were used for training, 20% for validation, and 10% for testing. Horizontal and vertical flips have been performed to augment. Along with the augmentations, pre-processing methods including normalization and standardization were used; as part of the process, all photos were treated to have a zero mean and unit standard deviation.

4.2 Metrics

1. *Loss*: Categorical cross-entropy has been chosen as the loss function for the implementation. When employing Softmax activation to solve multiclass classification issues, this loss function is extremely well-liked.

$$\text{Categorical cross-entropy} = n \sum_{\text{i=1}} t_i . \log(p_i)$$

2. *Accuracy*: The model's accuracy is determined by comparing its prediction to the actual data. The total accuracy is evaluated as

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

3. *Precision*: The percentage of true positives chosen from the entire number of positive instances, or what proportion of the positive data is truly forecasted as positive, is how a model's recall is determined.

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

4. *Recall*: The percentage of true positives chosen from the entire number of positive instances, or what proportion of the positive data is truly forecasted as positive, is how a model's recall is determined.

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

4.3 Training Details

The TensorFlow framework was used to code the SLRNet architecture, and the NVIDIA P100 GPU and CUDA integration were combined to train it. With a batch size of 32, 38,500 photos were used to train the model. Before being loaded into a data loader object for training, each picture underwent pre-processing and was enhanced, normalized, and standardized. To reduce the category cross-entropy loss, the model was trained using the Adam optimizer for 10 iterations at a learning rate of 0.001.

4.4 Experimental Details

With a batch size of 32, validation was done on 11,000 photos using the NVIDIA P100 GPU. The weights were hosted, downloaded, and used for additional tests. Figure 4 displays testing outputs from SLRNetwork, which correctly classified the input images into their individual classes with a validation accuracy: 0.9908, precision: 0.9938, and recall: 0.9912 (Figs. 3 and 5).

Fig. 3 Test results

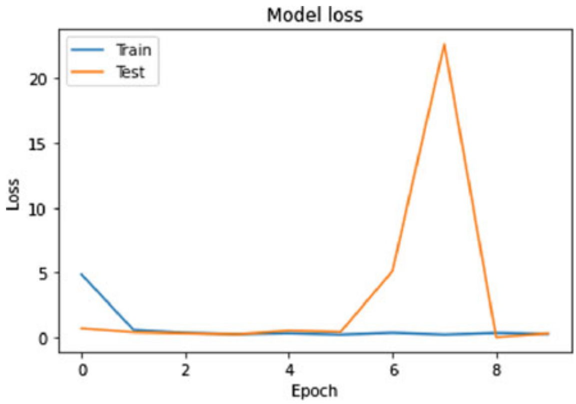


Fig. 4 Loss progression

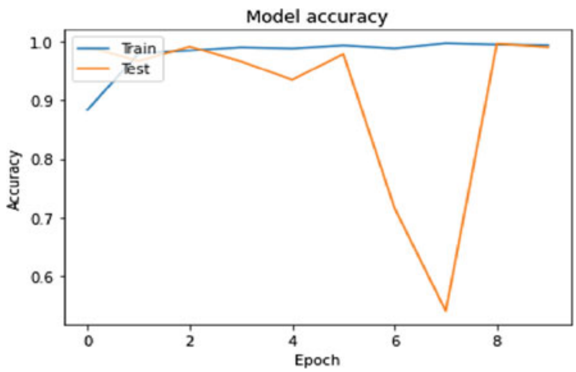


Fig. 5 Accuracy progression

5 Conclusion and Future Works

All individuals who previously lacked the capacity to express their opinions to society may now speak with one another via sign language. This sector will be significantly impacted by the creation of an accurate sign language recognition system with no technological issues. In our study, we have attempted to retain high accuracy while removing one such flaw seen in the earlier models used for sign language recognition. The research project may be furthered by teaching the network to detect sentences and phrases rather than only letters, by including pertinent image data and labels within the data set, and by appropriately augmenting them. Deploying and automating the process flow, will further expand the research scope.

References

1. Kasapbaşi A, Ahmed Elbushra AE, Al-Hardanee O, Yilmaz A (2022) DeepASLR: a CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals. Comput Methods Programs Biomed 2:100048
2. Koller O, Ney H, Bowden R (2016) Deep hand: how to train a CNN on 1 million hand images when your data is continuous and weakly labeled. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3793–3802
3. Koller O, Zargaran O, Ney H, Bowden R (2016) Deep sign: hybrid CNN-hmm for continuous sign language recognition. In: Proceedings of the British machine vision conference
4. Cui R, Liu H, Zhang C (2017) Recurrent convolutional neural networks for continuous sign language recognition by staged optimization. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7361–7369
5. Pu J, Zhou W, Li H (2019) Iterative alignment network for continuous sign language recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4165–4174
6. Ferriera P (2018) Multi-modal learning for sign language recognition. In: Pria National Conference
7. Huang J, Zhou W, Zhang Q, Li H, Li W (2018) Video-based sign language recognition without temporal segmentation. In: Thirty-second AAAI conference on artificial intelligence
8. Quan Y, Jinye P (2008) Chinese sign language recognition for a vision-based multi-features classifier. In: International symposium on computer science and computational technology, pp 194–197. https://doi.org/10.1109/ISCSCT.2008.374
9. Athitsos V, Wang H, Stefan A (2010) A database-based framework for gesture recognition. In: Personal and ubiquitous computing, pp 511–526
10. Rana S, Liu W, Lazarescu M, Venkatesh S (2009) A unified tensor framework for face recognition. Pattern Recogn 42(11):2850–2862
11. Flasiński M, Myśliński S (2010) On the use of graph parsing for recognition of isolated hand postures of polish sign language. Pattern Recogn 43(6):2249–2264
12. Karami A, Zanj B, Sarkaleh AK (2011) Persian sign language (PSL) recognition using wavelet transform and neural networks. Expert Syst Appl 38:2661–2667

Social Harassments Cyberbullying Detection Using Machine Learning

Bhavana Jamalpur, D Kothandaraman, A. Navya, G. Bhavya Keerthi, D. Srijanjali, and M. Sushmitha

Abstract Because of the widespread use of online social networks and the high quality of different online platforms, we now have more opportunities than ever before. Unfortunately, most teenagers are victims of cyberbullying. Most of students in online are embarrassed, bullied, and harassed by unknown users and strangers through the use of posting of negative words, harmful, false, or mean content regarding. This can result in stress, unhappiness, depression, and even suicidal ideation. The main purpose of this research is to pinpoint and categorize offensive and non-offensive words used in social media posts as either harassment words/ speech or not. Machine learning classifiers were used to determine whether tweets were offensive or not.

Keywords Classifiers · Offensive · Cyberbullying · Bullied · Tweets

1 Introduction

As social media like Instagram, Snapchat, and TikTok have gained popularity, youth are passing their most of the time online navigating a complex virtual world. In cyberbullying, people use technology to send threatening or awkward messages to another individuals. Cyberbullying or social media bullying incorporates activities and measures to regulate, get at, or stigmatize anyone. These terrible activities are solemnly harming and might influence anybody effectively and seriously. They primarily happen via web-based media, public gatherings, and alternative online sites. At the suitable time, cyberbullying comes in numerous varied structures. It does not extremely mean hacking somebody's profile or presenting to be the person whom you already known. Bullying on social media is even worse because of its

B. Jamalpur (✉) · D. Kothandaraman · A. Navya · G. B. Keerthi · D. Srijanjali · M. Sushmitha
School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India
e-mail: j.bhavana@sru.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_25

fast spread to the broader audience. Analysis reveals that such negative behavior most often occurred on Facebook/Twitter [1]. Of the 80.83 million Internet users [2] across the world, over ninety percentage of social media people use Twitter, mostly young, vulnerable, and vulnerable. Lot of material and the data available online sites are generally unprocessed, dirty, noisy, and usually a combination of many skill semantic languages, and analysis shows the use of machine learning algorithms [3] to recommendation system [4] on knowledge of online media which has more correctness than the knowledge gained through social media. Search, Keywords and Content Analysis [5, 6].

2 Related Work

This study focused on cyber bullying on Twitter/tweets. Identifying different abusers on Twitter movements involved in the grouping the tweens and categorizing and labeling of information gathering the data from different sites and use models and feature using machine learning techniques [7]. The Twitter Live API is used to collect tweets and train the dataset. The model tests both support vector machines and Naive Bayes use datasets to extract and retrieve features using the TF-IDF victimizers [2]. A novel approach to detect cyberbullying on different sites media platforms using a BERTmodel using "single linear neural network layer (SNN)" as a classifier on top [8]. Many applications such as "Mobicip" introduce total control such as category blocking, session timings, Internet usage, blocked expressions, and YouTube filtering, whereas others applications such as iAnon and GoGoStat search for specific profanity words [9, 10]. Some people use negative content such as emojis, mimes, and abbreviations that would be used in negative sense [11, 12].

A cyber-attack widely applied to text analysis approaches to online reviews/ comments. This approach yields higher accuracy and fewer false positives than simple list-based profanity word matching [13]. Naive Bayes classifier was used to retrieve inappropriate terms/words in Twitter text data to detect bullying [14]. Cyberbullying affecting the daily life of half of the American teens incidents in the Instagram social network, Instagram mime sand their comments. A model using a Naïve Bayes classifier was used to find hate words posted in Twitter [15]. Improving the accuracy of cyberbullying detection using Baye's classifier [16]. The problem of scalability is an important factor for other research areas such as mischievousness behavior detection in online video chat services online advertising and community-based cyber-attack detection [17, 18].

3 Problem Statement

This is a study of detecting the cyberbullying activities in social media like Twitter using various models of machine learning which indicates whether the tweets are offensive or non-offensive.

Online social network bullying is one of the common problems in social media sites. The development and advancement of technology in addition to having a positive impact also introduces new problems when used inappropriately or violates what they should, this is often referred to as cybercrime. Cyberbullying means harassing or insulting a person by using electronic communication through which victims can experience a wide range of emotions, with negative consequences and even ending their lives. Cyberbullying operations pose a serious threat to the victims' mental and physical health. Although a lot of research and the effects and nuisance of bullying is available, deployments for watching social networks to discovering of online harassment activities are unusual. As a result, the designed methodology focuses on detecting cyberbullying behavior in social networks.

4 Proposed Methodology

The tweets dataset contains real-world examples of tweets and messages scraped from online social media sites, offensive and non-toxic, i.e., non-offensive sentences. It also contains a diverse collection of negative words that are commonly used by people in their daily lives. This would aid us in detecting nearly every negative comment or tweet. The next step is to preprocess the data after it has been extracted. It is performed because real-world data contains a large number of unnecessary characters, necessitating data cleaning in order to prepare the data for the detection phase. This is a time-consuming but critical task (Figs. 1 and 2).

In this paper, initially collecting the various tweets in the form of the formatted dataset that stores the information that is collected from the tweets, comments, and hashtags or we can also have the already existing dataset from the Kaggle website, where the dataset consists of the required information for processing of the project, i.e., the dataset which also includes both the offensive and the non-offensive data.

In addition to that, the system by preprocessing of the data in the dataset which means the removal of the noisy data and the tokenization along with stemming of the data for the removal of the duplicated data in the dataset which helps in increasing of the accuracy and the decreasing if the time for the classification of the data, the system by training, and testing of the dataset, it will send the data for the cyberbullying detection. In the final step, the system classifies the data using support vector machines and random forest classifiers to classify data from tweets into offensive and non-offensive information. Along with the classification of the data, we can even calculate the precision and the accuracy of the data using the confusion matrix which the information of the data accurately.

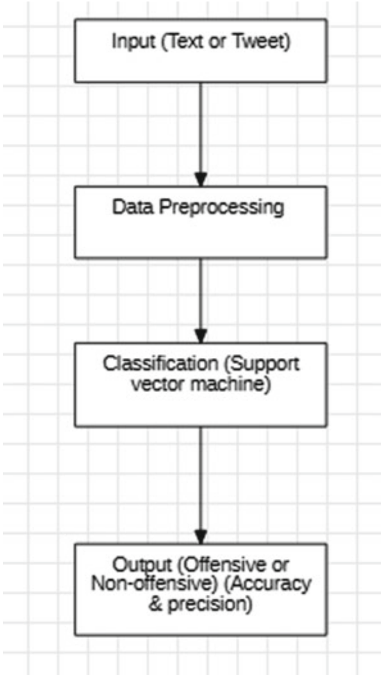


Fig. 1 Processing steps for detection of the cyberbullying

“Support vector machines (SVMs)” are used to explain regression classification problems. However, they are mainly used to solve classification problems. This project uses this classification method to classify data into offensive and non-offensive information.

5 Dataset

As the use of social media becomes more prevalent among all age groups, the majority of users depend on this media for their regular communication. The prevalence of media means that cyberbullying can cause harm to people for anyone at any time. Also, the relative privacy of the Internet makes such interpersonal outbreaks more difficult to thwart than bullying.. The information we gathered from Kaggle website provides three columns (Id, Label, and Tweets) that can be used to detect whether they are bullying or not. This dataset contains the 8799 rows and 3 columns in the Label which shows the tweets result like offensive and non-offensive.

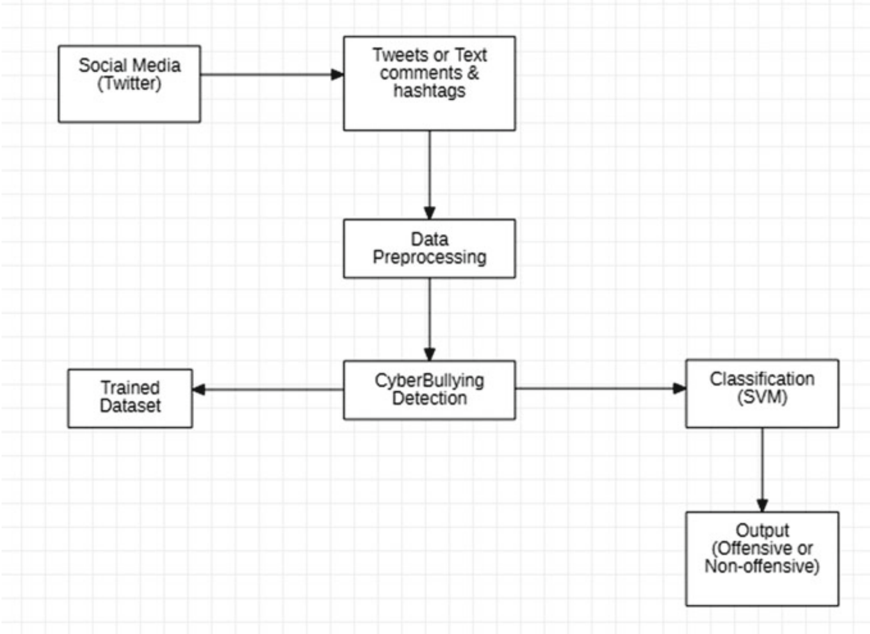


Fig. 2 Flowchart for the detection of the cyberbullying on Twitter

6 Result Analysis

In this code, sample will show the Label column from the dataset which declares the tweets as offensive and non-offensive. So, after writing code we will get the output in terms of graphs, and the purple graph represents offensive and light-red graph represents non-offensive (Fig. 3).

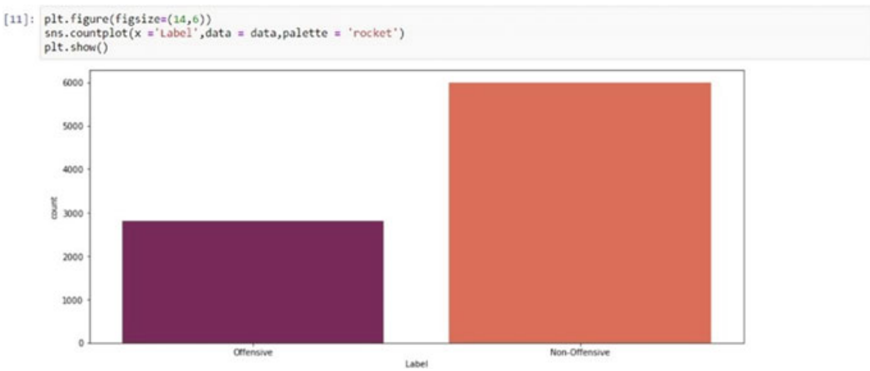


Fig. 3 Finding the offensive and non-offensive tweets

```
34]: plt.figure(figsize = (7,7))
     sorted_counts = data['Label'].value_counts()
     plt.pie(sorted_counts, labels = sorted_counts.index, startangle = 90, counterclock = False, wedgeprops = {'width' : 0.6},
             autopct='%1.1f%%', pctdistance = 0.7, textprops = {'color': 'black', 'fontsize' : 15}, shadow = True,
              colors = sns.color_palette("Paired")[7:])
     plt.text(x = -0.35, y = 0, s = 'Total Tweets: {}'.format(data.shape[0]))
     plt.title('Distribution of Tweets in the Dataset', fontsize = 16);
```

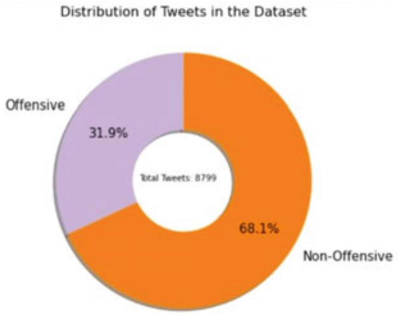


Fig. 4 Finding the percentage of offensive and non-offensive tweets

```
31]: acc1 = accuracy_score(y_test, y_pred)

32]: print(f"Accuracy of Support Vector Machine (Using Bag of words technique): {acc1}")

     Accuracy of Support Vector Machine (Using Bag of words technique): 0.007954545454545454
```

Fig. 5 Accuracy of the dataset

In the below code, sample we will show the distribution of tweets using pie-chart. After writing the code, we will get the output in terms of percentages, and the mauve color represents offensive and orange color represents non-offensive (Fig. 4).

After training the model and calculating confusion matrix, this have validated the accuracy by training the model. By using SVM model, we will find the accuracy of the dataset (Fig. 5).

In the code sample, we will show offensive tweets which we have taken from the dataset. After executing the code, we will get offensive words as output. Figure 6 shows some offensive words like stupid, fuck, and bitch, etc. (Fig. 7).

```
cloud1=WordCloud(background_color='black',colormap="Dark2",collocations=False,width=2000,height=1000).generate(" ".join(te

plt.axis('off')
plt.title("Offensive",fontsize=40)
plt.imshow(cloud1)
```

```
<matplotlib.image.AxesImage at 0x7faf17258210>
```

Fig. 6 Offensive tweets

```
cloud1=WordCloud(background_color='black',colormap="Dark2",collocations=False,width=3000,height=2000).generate(" ".join(text_Nonoffensive))

plt.axis('off')
plt.title("NonOffensive",fontsize=40)
plt.imshow(cloud1)
```

```
<matplotlib.image.AxesImage at 0x7fa14c487950>
```

Fig. 7 Non-offensive tweets

7 Conclusion

While the use of social networks and heavy use of the Internet have obvious benefits for society, their heavy use can also have significant negative effects which consists of sexual exposure, pranking result in crime, and bullying. A model is designed to detect cyberbullying activity and their misbehavior and its severity on Twitter. The system that we have used in this project will help in making the people by having the exact information about the required data which gives the baseline for separating the text and to classify the data. The developed model is a feature-based model that uses features of tweet content to classify tweets as offensive or non-offensive.

References

1. Patchin JW (2019) Summary of our cyberbullying research (2007–2019), Cyberbullying Research Centre

2. Zhao R, Mao K (2015) CyberBullying detection based on semantic-enhance marginalize denoising autoencoders. IEEE Trans Affect Comput
3. Andriansyah M, Akbar A, Ahwan A, Ariesto Gilani N, Roma Nugraha A, Nofita Sari R, Senjaya R (2017) Cyberbullying comment classification on Indonesian Selebgram using support vector machine method. In: Informatics and computing (ICIC) 2017 second international conference on, pp 1–5
4. Dinakar K, Reichart R, Lieberman H (2011) Modeling the detection of textual cyberbullying. In: The social mobile web
5. Ketsbaia L, Issac B, Chen X (2020) Detection of hate tweets using machine learning and deep learning. In: 2020 IEEE 19th international conference on Trust security and privacy in computing and communications (TrustCom), pp 751–758
6. Ali Shaik M, Verma D (2020) Deep learning time series to forecast COVID-19 active cases in INDIA: a comparative study. In: 2020 IOP conference series: materials science and engineering 981:022041
7. Sharma C, Ramakrishnan R, Pendse A, Chimurkar P, Talele KT (2021) Cyber-bullying detection via text mining and machine learning. In: 2021 12th International conference, computing communication and networking technologies (ICCCNT), pp 1–6
8. Kumar VN, Binsu C, Kovoor, Sreeja MU (2018) Cyber-bullying revelation in twitter data using Naïve-Bayes classifier algorithm. Int J Adv Res Comput Sci 9
9. Kothandaraman D, Balasundaram A, Korra SN, Sudarshan E, Vijaykumar B (2020) Enhancing dull images using discrete wavelet families and fuzzy. In: IOP conference series: materials science and engineering, vol 981, no 2
10. Pranathi V, Ranadheer Reddy G, Sunil G, Raghava Kumar D, Jamalpur B (2020) A comprehensive study on the various applications of deep learning. In: IOP conference series: materials science and engineering, vol 981, no 2
11. Salauddin MD, Nawaz Pasha S, Ramesh D, Kothandaraman D ()2019 An research to enhance the old manuscript resolution using deep learning mechanism. Int J Innov Technol Explor Eng 8(6s4)
12. Shaik MA, Verma D, Praveen P, Ranganath K, Yadav BP (2020) RNN based prediction of spatiotemporal data mining. In: IOP Conference Series: Materials Science and Engineering, vol 981, p 022027
13. Balasundaram A, Kothandaraman D, Ashokkumar S, Sudarshan E (2022) Chest X-ray image based COVID prediction using machine learning. In: AIP conference proceedings, vol 2418, issue 1, pp 020079
14. Ptaszynski M, Dybala P, Matsuba T, Masui F, Rzepka R, Araki K, Momouchi Y (2010) In the service of online order tackling cyberbullying with machine learning and affect analysis
15. Jamalpur B, RajyaLaxmi M, Kothandaraman D, Sudheer Kumar K (2022) Opinion mining on restaurant rating based on aspects. In: Data engineering and intelligent computing, pp 489–496
16. Kowalski RM, Limber S, Limber SP, Agatston PW (2012) Cyberbullying: bullying in the digital age. John Wiley & Sons
17. Chavan VS, Shylaja S (2015) Machine learning approach for detection of cyber-aggressive comments by peers on social media network. In: 2015 International conference on IEEE advances in computing, communications and informatics (ICACCI), pp 2354–2358
18. Sanchez H, Kumar S (2012) Twitter bullying detection. In: NSDI 2012, Berkeley, CA, USA, p. 15. USENIX Association (2012)

Smart Parking System Using Raspberry Pi

Bura Vijay Kumar, Khaja Mannan, Mothe Rajesh, D. Kothandaraman, A. Harshavardhan, and P. Kumaraswamy

Abstract This paper presents a new method for the people who want to park their vehicle in the modern city, and also an algorithm was implemented for the current cloud-based smart parking system using Raspberry Pi dependent on the Internet of Things innovation. The framework that helps people to go through the website and he can check the free slots to park the vehicle and the client can also check complete number of free places; with this smart parking system, the client can save his time, fuel, and also he can relief from the mental tension to park his vehicle. Parking vehicle issue became one of significant tasks in the city transportation the executives since the relating to or occupying space asset of a city is restricted and the stopping cost is costly. Bunches of vehicles out and about ought to invest superfluous time and devour energy during looking for leaving because of restricted parking spot. To adapt to these constraints and give more savvy answers for drivers in the determination of stopping office, this examination proposes a brilliant stopping direction calculation. The proposed calculation upholds drivers to track down the most suitable stopping office thinking about constant status of stopping offices in a city. To recommend the most appropriate stopping office, a few factors like driving distance to the directed stopping office, strolling distance from the directed stopping office to objective, anticipated stopping costs, and gridlock because of stopping direction, are having been thought about carefully in the projected calculation. The projected calculation

B. V. Kumar (✉)
Center for Embedded Systems and Internet of Things, School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India
e-mail: vijaykumar.bura@gmail.com

K. Mannan
Sumathi Reddy Institute of Technology for Women, Warangal, Telangana, India

M. Rajesh · D. Kothandaraman
School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India

A. Harshavardhan
VNR Vignana Jyothi Institute of Technology and Engineering, Secunderabad, Telangana, India

P. Kumaraswamy
Kakatiya Institute of Technology and Science, Warangal, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_26

assists with augmenting the use of room assets of a metropolitan city, and diminishes superfluous energy utilization and carbon dioxide (CO_2) emanation of meandering vehicles since it is intended to manage the usage of leaving office effectively and lessen gridlock because of parking spot explore.

Keywords Smart parking system (SPS) · City transportation management · Smart parking guidance · Parking guidance algorithm

1 Introduction

In the improvement of traffic, the board frameworks, an insightful leaving framework was made to diminish the expense of employing individuals and for ideal utilization of assets for vehicle leave proprietors. Right now, the normal strategy for discovering a parking spot is manual, where the driver for the most part discovers a space in the road through fate and experience. This cycle requires some investment and exertion and may prompt the most pessimistic scenario of neglecting to discover any park space if the driver is driving in a city with high vehicle thickness. The option is toward discovering a predefined vehicle leave through high limit. Nonetheless, this is certifiably not an ideal arrangement in light of the fact that the vehicle park could typically be far away from the client objective. Past study has utilized vehicle-to-vehicle (V to V) [1] and vehicle-to-foundation (V to F) [2] connection with the help of different remote organization advancements, e.g., radio recurrence recognizable proof (radio frequency identification), Zigbee, remote wreck network [1], and the Internet. This examination meant to give data about close by parking spots for the driver and to reserve a spot minutes sooner utilizing upheld gadgets, e.g., cell phones or tablet personal computers. Besides, the administrator identifies the users and helps the users to book their parking spot. Notwithstanding, the present savvy leaving framework doesn't give an in general ideal arrangement in tracking down an accessible parking spot, doesn't take care of the issue of burden adjusting, doesn't give monetary advantage, and doesn't get ready for vehicle-refusal administration. To determine the previously mentioned issues and exploit the huge advancement in innovation, the Internet of Things (IoT) innovation has made a transformation in numerous fields in life just as in brilliant stopping framework smart parking system innovation [1]. The current investigation proposes and fosters a powerful cloud put together smart parking system arrangement based with respect to the Internet of Things. Our framework builds every vehicle leave as an Internet of Things organization, and the information that incorporates the motor vehicle GPS area, distance between vehicle leaving regions, and number of free slots in vehicle leave regions will be moved to the server.

The cloud server helps to determine the free parking slots and charge for the parking based on time stamp, and these free parking slots and the parking expenditure are continuously reorganized and are available any cost of time by the motor vehicle in the system. The smart parking system is a new technology, with this technology we can automatically supervise and manage the car parking system and each car parking system can function separately as a traditional smart car parking system.

Initially, a car (from this point forward a vehicle) was developed to expand accommodation and support in regular day to day existence. Even so, the vehicle clog in a major city messes unfortunate up like natural issues, energy utilization, parking spot deficiency, gridlocks, clamor, air contamination, and surprisingly minor mental harm to certain individuals. Due to huge traffic and issues like lack of parking spots and lack of prior knowledge on parking system. Thus, numerous vehicles out and about ought to invest pointless time and burn-through unessential energy during looking for parking spots. As indicated by the new examination work [3] managing the meaning of stopping issue, the traffic stream top brought about via looking through stopping offices can increment as much as around 25–40%. Chegini et al. [4] referenced that about 30% of vehicles out and about in the midtown space of significant urban communities appeared to journey for parking spaces, which took a normal of 7.8 min. The other examination tracked down that the meandering of vehicles to discover a leaving office is liable for about 30% of the whole traffic in a city [5] referred to from an investigation of stopping circumstance in Schwa Bing (an area of Germany) that a yearly complete economy harm had been assessed as 20 million euro, caused simply by the traffic looking with the expectation of complimentary parking garages [6, 7].

2 Related Work

In certain examinations, the creators anticipated a new algorithm for smart parking system for vehicles. In the first part, they utilized a calculation to plan the online issue of a stopping framework into a disconnected issue. In second part, they set up a numerical model depicting the disconnected issue as a straight issue. In third part, they planned a calculation to tackle this straight issue. At long last, they assessed the proposed calculation utilizing exploratory re-enactments of the framework. The test results demonstrated opportune and effective execution. Nonetheless, these papers don't specify the asset reservation instrument (all leaving prerequisites are inferred promptly and are put in the line), the component for surveying the assets framework, the instrument to direct vehicles to the parking spot, the system for taking care of circumstances when the solicitation for administration is denied and don't figure the normal holding up time and normal complete time that every motor vehicle spend on the framework [3]. In another examination, they propose a smart parking system dependent on the combination of UHF recurrence, wireless sensor network innovations radio frequency identification (RFID), and IEEE 802.15.4. This framework can gather data about the condition of inhabitance of the vehicle leaves and can guide drivers to the closest empty parking space by utilizing a product application. Be that as it may, in this work, the creators have no numerical conditions for the framework design and don't make a huge scope stopping framework [8]. The consequences of this paper just carry out the proposed engineering; they don't make reference to the presentation of the smart parking system framework [5, 9]. Proposed an imaginative framework including the stopping direction administration. The parking spot held by a cell phone through Internet access. After entering the vehicle in the saved parking spot will be shown as filled with vehicle in red color as shown in smart parking framework. An inertial route framework is carried out to direct the vehicle to the held space. The framework will occasionally refresh the situation with the parking spot continuously to assist with guaranteeing framework precision. Framework execution is estimated through the exactness of the inertial route frameworks run in an indoor climate, and the framework execution is assessed by thinking about the precision of the global positioning system. In this research, the creators have not assessed the presentation of the leaving administrations; they don't give any numerical model of the framework, and don't think about the hanging tight season of every vehicle for administration. Different specialists have planned engineering for stopping the executives in brilliant urban areas [6, 7].

They proposed smart stopping collaborator (IPA) engineering pointed toward defeating current public stopping the board arrangements. This engineering gives drivers data about on-road stopping slow down accessibility and permits drivers to save the most advantageous stopping slow down at their objective before their takeoff. They use radio frequency ID innovation in this framework. At the point when no a vehicle stops or leaves the IPA parking space, the radio frequency ID authorized person and the attractive circle distinguish the activity and send this data to the unit regulator to refresh the data on the vehicle leave status [10, 11]. This examination utilizes just some straightforward numerical conditions for the framework design and doesn't make a large-scale stopping framework. In different works, creators have planned and carried out a smart parking system [12, 13] to take care of the stopping issue. A piece of this framework is carried out in the Zigbee network which sends pressing data to a personal computer through an organizer and afterward refreshes the dataset. The application layer can rapidly ignore the stopping data through the Internet and utilize the upsides of a web administration to accumulate all the dispersed stopping data for the accommodation of the individuals who need to discover a parking spot [9, 14].

The paper essentially reports the plan and execution of a smart parking system and doesn't assess the framework execution [7]. Expected to mechanize the vehicle and the vehicle leaving. This paper examines a venture which present a smaller than normal model of a computerized vehicle leaving framework that can direct and deal with the quantity of vehicles that can be left in a given region at some random time dependent on the accessibility of parking spots. The mechanized leaving strategy permits the leaving and leaving of vehicles utilizing detecting gadgets. Section to or exit from the vehicle leave is instructed by a web-based application. The contrast between the Bonde framework and the other existing frameworks is that the creators were planning to make the framework as minimal human reliant as conceivable via computerizing the vehicles just as the whole vehicle leave; then again, most existing frameworks require human intercession (the vehicle proprietor or other) to leave the vehicle [9, 15], which portrayed another smart parking system design dependent on the Internet of Things innovation. The design of this framework comprises a wireless sensor network (WSN) and Zigbee which is a front-end layer and IoT middleware layer as the internal user interface that gives information answering to the client. Be that as it may, there is detriment as it doesn't utilize an appropriate application convention for the exchange of information from the wireless sensor network to the worker, like the constrained application protocol (CoAP). There are no framework for mathematical and execution assessment [16, 17].

3 Implementation

Smart parking system framework first recognizes the vehicle left in quite a while and shows the situation with each opening through the web server. The people who want to park their vehicles can know the situation with the leaving opening through the website page, after that they need to show their radio frequency identification (RFID) to the authorized person (to read or examine with care) to get confirmed and note down the hour of that particular vehicle. Assuming the new guest is identified through another radio frequency identification (RFID) authorized person, it prompts to ask the username, secret phrase, and mail ID with their radio frequency identification (RFID) and store those subtleties to the Raspberry Pi. Then, at that point while leaving, by showing the frequency identification (RFID) to the authorized person, it notes down the out time and computes the complete time with the cash for that time. Then, at that point, it will be identified from the record balance and the implication will be shipped off their personal mail.

4 Components Required

See Fig. 1.

5 Circuit Diagram

See Fig. 2.

5.1 Flowchart

See Fig. 3.

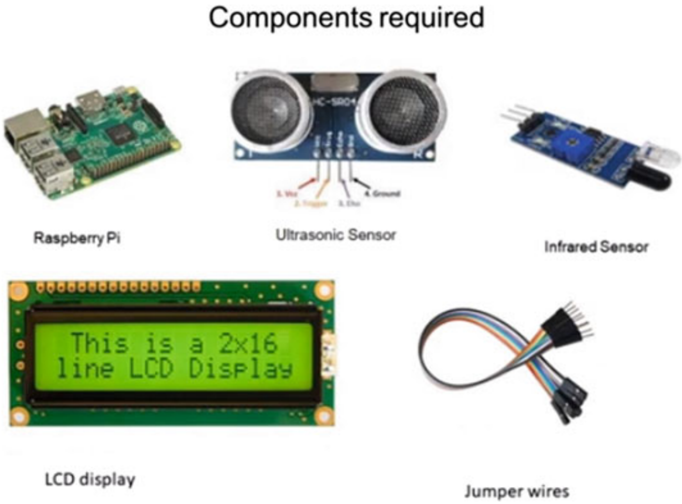


Fig. 1 Components required developing the system

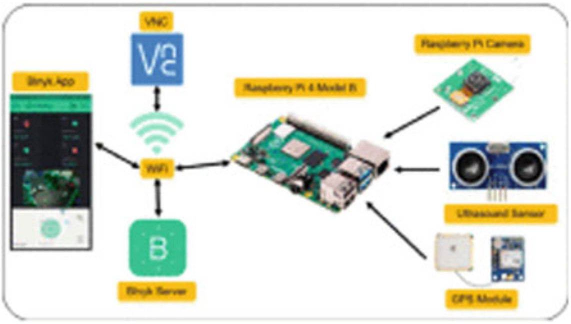
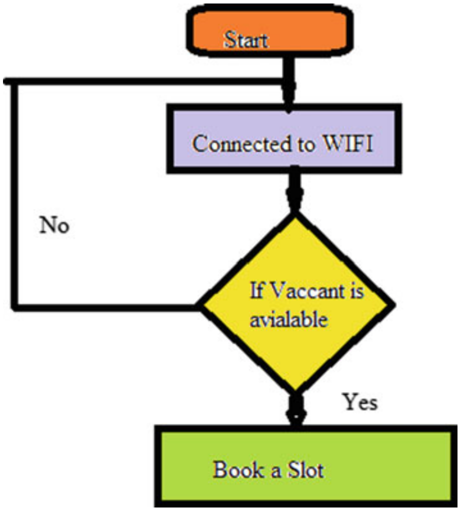


Fig. 2 Circuit design for developing the system

Fig. 3 Flowchart



6 Results Analysis

Slots	Slot 1	Slot 2	Slot 3	Slot 4
Vehicle	Filled by Vehicle	Filled by Vehicle	Filled by Vehicle	Filled by Vehicle

All the slots are filled with Vehicles,
No Empty Slot to park the vehicle

Slots	Slot 1	Slot 2	Slot 3	Slot 4
Vehicle	Empty	Filled by Vehicle	Filled by Vehicle	Filled by Vehicle

Slot 1 is empty, means a user can park the vehicle
slot 2,3 and 4 are filled with vehicles

Slots	Slot 1	Slot 2	Slot 3	Slot 4
Vehicle	Empty	Empty	Empty	Empty

Slot 1,2,3 and 4 are empty, means a users can park the vehicles

Slots	Slot 1	Slot 2	Slot 3	Slot 4
Vehicle	Empty	Filled by Vehicle	Empty	Filled by Vehicle

Slot 1,3 are empty, means a user can park the vehicles in 1 and 3 slots.
slot 2 and 4 are filled with vehicles

7 Conclusion

The parking framework turns out to be further developed than before which has the information base, mail notices, fees assessment, auto-identification, and so on.

References

1. Shafique K, Khawaja BA, Sabir F, Qazi S, Mustaqim M (2020) Internet of Things (IoT) for next-generation smart systems: a review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. IEEE Access 8:23022–23040
2. Elakya R, Seth J, Ashritha P, Namith R (2019) Smart parking system using IoT. Int J Eng Adv Technol (IJEAT) 9(1):6091–6095
3. Madakam S, Ramaswamy R, Tripathi S (2015) Internet of Things (IoT) a literature review. J Comput Commun 03:164–173
4. Chegini H, Naha RK, Mahanti A, Thulasiraman P (2021) Process automation in an IoT–fog–cloud ecosystem: a survey and taxonomy. IoT 2:92–118
5. Farooq UM et al (2015) A review on internet of things (IoT). Int J Comput Appl 113(1):1–7
6. Ammar M, Russello G, Crispo B (2018) Internet of Things: a survey on the security of IoT frameworks. J Inf Secur Appl 38:8–27
7. Nguyen-An H et al (2021) IoT traffic: modeling and measurement experiments. IoT 2(1):140–162
8. Swamy SN, Kota SR (2020) An empirical study on system level aspects of Internet of Things (IoT). IEEE Access 8:188082–188134
9. Biswas AR, Giaffreda R (2014) IoT and cloud convergence: opportunities and challenges. In: 2014 IEEE world forum on Internet of Things (WF-IoT). IEEE
10. Tawalbeh L et al (2020) IoT privacy and security: challenges and solutions. Appl Sci 10(12):4102
11. Kothandaraman D, Manickam M, Balasundaram A, Pradeep D, Arulmurugan A et al (2022) Decentralized link failure prevention routing (DLFPR) algorithm for efficient internet of things. Intell Autom Soft Comput 34(1):655–666
12. Ngu AH et al (2016) IoT middleware: a survey on issues and enabling technologies. IEEE Internet Things J 4(1):1–20
13. Kothandaraman D, Balasundaram A, Dhanalakshmi R, Sivaraman AK, Ashokkumar S et al (2022) Energy and bandwidth based link stability routing algorithm for IoT. Comput Mater Continua 70(2):3875–3890
14. Chettri L, Bera R (2019) A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. IEEE Internet Things J 7(1):16–32
15. Hwang YH (2015) IoT security & privacy: threats and challenges. In: Proceedings of the 1st ACM workshop on IoT privacy, trust, and security
16. Stergiou C et al (2018) Secure integration of IoT and cloud computing. Future Gen Comput Syst 78:964–975
17. Nizetic S et al (2020) Internet of Things (IoT): opportunities, issues and challenges towards a smart and sustainable future. J Cleaner Prod 274:122877

Home Devices Controlling and Monitor System Using IoT

Bura Vijay Kumar, D. Kothandaraman, N. Varsha, T. Soumya Sri, A. Sumanth, G. Sruthi, and D. Rakesh Singh

Abstract The suggested system primarily intends to create a fully automated home using a touch screen to manage household appliances and to provide a user-friendly environment for the user to successfully operate the equipment. This solution allows you to operate your household appliances from any location on the planet. The system includes a touch control panel that is meant to monitor electrical equipment such as lights, fans, and coolers, allowing us to turn them off using the touch control panel. This technology creates an environment in which the user may operate the household appliances by tapping the desired places on the touch screen. This is a system that assists a user in using numerous switches gadgets for lighting and appliances. The input device, a touch screen, is significantly easier to use. Touch screens are usually regarded as the most user-friendly input device available. In a smart home context, this system will use home network and cloud computing technologies to operate and monitor home equipment (fan, light, and chiller).

Keywords Smart home automation system · Cloud computing technologies · Central smart home hub

B. Vijay Kumar (✉)
Center for Embedded Systems and Internet of Things, School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India
e-mail: vijaykumar.bura@gmail.com

D. Kothandaraman
School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India

N. Varsha · T. Soumya Sri · A. Sumanth · G. Sruthi · D. Rakesh Singh
Computer Science and Engineering, SR Engineering College, Warangal, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_27

1 Introduction

The smart home automation system allows you to operate your home appliances from anywhere in the globe. This system offers an environment in which the user may operate the household appliances by giving orders by tapping the desired places on the touch screen. This home automation system is utilized to make our surroundings safer, more energy efficient, and to save our important time. Rapid technological advancements are reducing the need for human labor [1, 2]. One of the technologies used to monitor household appliances is home automation. This automation system does not require any mobile phones to manage the household appliances since we cannot monitor the gadgets if the phone goes off. The touch control panel is intended to monitor electrical equipment [3, 4].

In most situations, a home automation system links controlled devices to a central smart home hub (sometimes called a "gateway"). The system's user interface consists of wall-mounted terminals, tablet or desktop computers, a mobile phone application, or a Web interface that is also available off-site through the Internet. Open-source solutions are growing increasingly popular, despite the fact that there are various providers to select from. The current state of home automation, however, has certain drawbacks, such as the lack of standardized security measures and the deprecation of older devices that are not backwards compatible. Home automation has a lot of potential for data sharing among family members or trusted persons for personal security, and it might eventually lead to energy-saving measures with a good environmental impact [5, 6].

2 Motivation and Scope of Work

This is our primary reason for creating a smart home automation system. Initially, home automation was created with the help of a mobile application. We have downsides with this system, such as the requirement to carry a cell phone to operate household equipment. We have built a system with a touch screen that controls all the appliances in the home to address this difficulty. We have designed a system with a touch screen that consumes more battery and may hang during operation. The goal of the project is to use technology to offer a comfortable living for human beings. As technology advances, dwellings become smarter as well. Humans will have a better quality of life as a result of the initiative [7, 8].

The ease of smart home solutions is what propels them forward. In today's world, convenience makes an important role which is called as "time saver," each and every second is precious, e.g., phones, on the other hand, get us information from other people faster; and computers, on the other hand, get us work done faster. Smaller home comforts will appeal to users since they enable the house to save time. Dishwashers, washing machines, and microwave ovens are just a few examples of domestic technology. These technologies are primarily mechanical in nature, and electrical household amenities are typically absent [9, 10].

The goal of a smart home system is to bring digital technology's benefits. A smart home system, for example, eliminates the need for the user to roam about turning off lights; instead, they may save time by just pressing a button on the touch pad or configuring the lights to turn off after a certain amount of time. Smart home systems will allow users to play music from computer on their sound system without having to go to computer, search for the song, and the song plays [10, 11].

3 Problem Statement

In our everyday busy schedule, we forget to control and monitor our home devices and we leave to our work which causes us power wastage. So, to maintain and control these devices from anywhere in the world, we use Node MCU, an open-source IoT platform, to design, develop, and simulate an IoT-based home automation system that allows complete control of all loads in the system from any location with an Internet connection. Home automation allows us to use a touch pad to operate equipment in our houses from anywhere on the earth. "Home automation" is more correctly linked with houses in which practically everything, including lighting, fans, heating, cooling, and ventilation systems, is connected to a remotely controllable network. We design a prototype that may be used to control and monitor household appliance fans and lights in this project [12, 13].

4 Literature Review

"Bluetooth-based home automation" is a concept that allows any smartphone to control it. The automation system interfaces with the smartphone through Bluetooth. The Bluetooth interface on an Android phone is used to send control signals to switch on or off household equipment [14]. "Android-based home automation" uses a simple Android app to control electrical devices with simple clicks or voice commands. Bluetooth is used to send commands to the Arduino Uno. As a consequence, you will not have to interrupt your movie or work to turn the device on or off [15]. "Android-based home automation with Raspberry Pi" reduces the requirement for human labor in the manufacturing of goods and services. Automation plays a vital influence in both the global economy and everyday lives. This article describes the design and construction of a remote household appliance control system utilizing a Raspberry Pi and an Android smartphone. You will be able to control every device in your house from anywhere on the planet once you have completed this Raspberry Pi home automation. The ability to operate the house from any location is critical [16]. "Home automation with artificial intelligence," with the introduction of AI to home automation, the user's living area can be controlled remotely or centrally without the

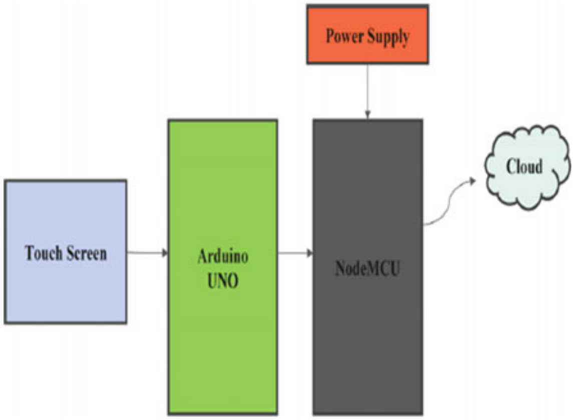


Fig. 1 Block diagram of transmitter

touch of a button. The required software tools should have a user-friendly interface. It is expected in the suggested system that the home system adjusts to the occupant's lifestyle. The proposed intelligent system is constantly learning and adapting to the preferences of the users [17, 18].

5 Project Description

5.1 *Transmitter*

This is the block diagram of transmitter shown in Fig. 1. The hardware components are NodeMCU, Arduino Uno, and TFT touch screen display.

5.2 *Receiver*

Figure 2 depicts the block diagram of a receiver. The hardware consists of a NodeMCU and household appliances like a fan and an LED.

- Because it has 16 analog pins, the touch screen is interfaced with the Arduino. Because a touch screen requires additional analog pins, it should be interfaced with the Arduino rather than the NodeMCU.
- Here, two NodeMCUs are used. They are NodeMCU at transmitter side and NodeMCU at the receiver side.
- Cloud computing is used for communication between the transmitter and the receiver.

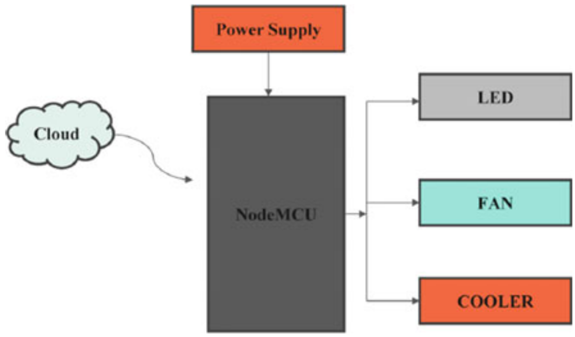


Fig. 2 Block diagram of receiver

- After interfacing the Arduino with touch screen, it should be connected to the NodeMCU in the transmitter side which consists of inbuilt Wi-Fi module.
- Data which is given to the NodeMCU transmitter side will be transmitted to the cloud, and NodeMCU at receiver side will receive the data from the cloud.
- After receiving the data, using touch screen, we can automatically control the home appliances like fan, light, and cooler from anywhere.

5.3 Flowchart

The flowchart shows that how the project will be executed when the portion of the touch screen is pressed (Fig. 3).

5.4 Design of TFT Touch Screen

- The touch screen is linked to the Arduino UNO, which has 14 analog pins.
- Touch screen is designed with four different portions such as LIGHT-ON, LIGHT-OFF, FAN-ON, and FAN-OFF.
- Touching the LIGHT-ON section of the touch screen immediately turns on the light, while touching the LIGHT-OFF portion automatically turns it off.
- By touching on the FAN-ON portion, it automatically turns on the fan, and by touching on the FAN-OFFF portion, it automatically turns off the fan (Fig. 4).

5.5 Interfacing Transmitter

- Touch screen is interfaced with the Arduino UNO.
- Then, it is connected to the NodeMCU in the transmitter side. So that data is transmitted from touch to cloud.

Fig. 3 Flowchart

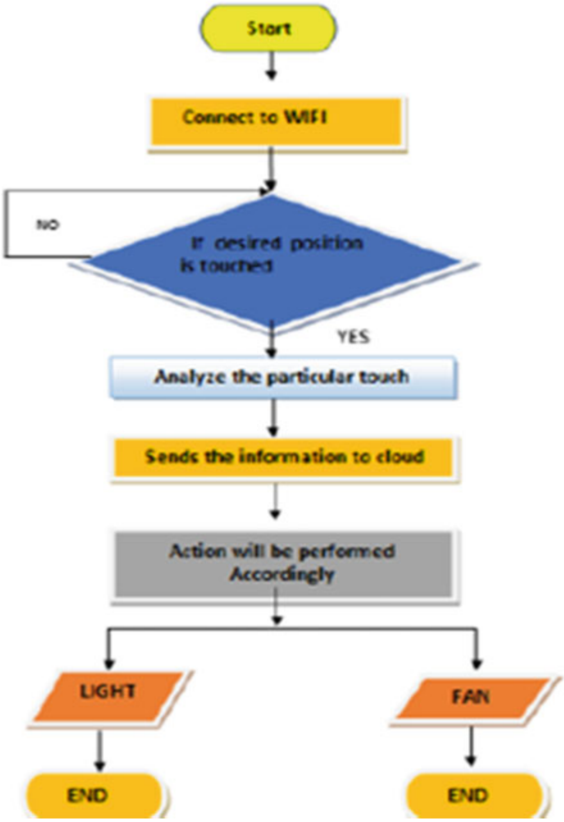


Fig. 4 TFT touch screen

- Touch pad is portable. We can control home appliances from anywhere (Fig. 5).

Fig. 5 Transmitter

Fig. 6 Receiver

5.6 Interfacing Receiver

- Data is received from the transmitter through cloud computing.
- Here, we have used another NodeMCU to receive the data. Home appliances can be controlled by receiving the data (Fig. 6).

6 Results and Discussion

See Figs. 7, 8, 9 and 10.

Fig. 7 When we press the LIGHT-ON button on the touch control panel, led light bulbs turn on

Fig. 8 When we press the LIGHT-OFF button, it stops glowing and the lights switch off

Fig. 9 When we press the FAN-ON button on the touch control panel, the fan turns on

Fig. 10 When we press the FAN-OFF button on the touch control panel, the fan goes off

7 Conclusion

Demotics is quickly becoming the most essential aspect in bringing an automated environment into every home. We can operate our household appliances from anywhere in the world thanks to this technology. We can make effective use of this technology without wasting time. This automated mode makes consumers' lives easier by completely automating critical appliances while still requiring some human work. Every human being's day-to-day life is drastically altered by this technology. Smart houses are gaining in popularity. This system's future potential includes making houses increasingly smarter. This system may be tightly connected with home security solutions to provide house owners more control and protection. The next stage is to expand this technology to automate large-scale environments like offices and factories. Smart houses can handle household appliances, lighting, the environment, energy management, and security, as well as connect to other networks, thanks to standardization.

References

1. Shafique K, Khawaja BA, Sabir F, Qazi S, Mustaqim M (2020) Internet of Things (IoT) for next-generation smart systems: a review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. IEEE Access 8:23022–23040
2. Madakam S, Ramaswamy R, Tripathi S (2015) Internet of Things (IoT): a literature review. J Comput Commun 03:164–173
3. Chettri L, Bera R (2019) A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. IEEE Internet Things J 7(1):16–32
4. Pramod Kumar P, Sandeep CH, Naresh Kumar S (2018) An overview of the factors affecting handovers and effective highlights of handover techniques for next generation wireless networks. Indian J Public Health Res Dev 9(11):722–725
5. Swamy SN, Kota SR (2020) An empirical study on system level aspects of Internet of Things (IoT). IEEE Access 8:188082–188134
6. Kumar BV, Chanti Y, Yamsani N, Aluvala S, Bhaskar B (2019) Design a cost optimum for 5g mobile cellular network footing on NFV and SDN. Int J Recent Technol Eng 8(2S3):1121–1129

7. Arabelli RR, Rajababu D (2019) Transformer optimal protection using internet of things. Int J Innov Technol Explor Eng 8(11):2169–2172
8. Vijay Kumar B, Chanti Y, Kothandaraman D, Harshavardhan A, Kanugula S (2019) Internet of things middleware architecture for communication. J Study Res 11(12):46–67
9. Ngu AH et al (2016) IoT middleware: a survey on issues and enabling technologies. IEEE Internet Things J 4(1):1–20
10. Kothandaraman D, Chellappan C, Sivasankar P, Pasha SN (2019) Context-aware energy conserving routing algorithm for internet of things. Int J Comput Netw Commun 11(3):15–32
11. Vijay Kumar B, Arabelli RR, Mothe R, Kothandaraman D, SeenaNaik K (2020) Smart traffic control system for vehicles on roads using Raspberry Pi. J Mech Continua Math Sci 15(9):162–170
12. Mohmmad S, Dadi R, Kothandaraman D, Sudarshan E, Pasha SN, Shaik MA (2022) A survey machine learning based object detections in an image. AIP Conf Proc 2418(1):020024
13. Kothandaraman D, Chellappan C (2019) Node rank based energy efficient routing algorithm for mobile Ad-hoc network. Int J Comput Netw Commun (IJCNC) 11:45–61
14. Kothandaraman D, Chellappan C (2017) Human activity detection system using internet of things. Int J Comput Sci Eng 9(11):657–665
15. Balasundaram A, Ashokkumar S, Kothandaraman D, Naikkora S, Sudarshan E, Harshaverdhan A (2020) Computer vision based fatigue detection using facial parameters. In: IOP conference series: materials science and engineering, vol 981, issue 2, pp 022005
16. Magesh Kumar S, AuxiliaOsvin Nancy V, Balasundaram A, Kothandaraman D, Sudarshan E (2020) Innovative task scheduling algorithm in cloud computing. In: IOP conference series: materials science and engineering, vol 981, issue 2, pp 022023
17. Chanti Y, Korra SN, Vijay Kumar B, Harshavardhan A, Kothandaraman D (2019) New technique using an IoT robot to oversight the smart domestic surroundings, StudiaRosenthaliana (J Study Res) 0039–3347
18. Chanti Y, Akarapu M, Swathi B, Vijaykumar B, Kothandaraman D (2022) Using the NI PXI platform for Li Fi-enabled intelligent transportation systems. AIP Conf Proc 2418(1):020071

Dynamic Text Generation from Sign Language—Review

Chekuri Sri Sumanth, D. Nanda Surendhar, G. Pramod, P. Pranav Rao, and T. Babu Karthik

Abstract Recognizing the sign language is very important in research fields because that is the only way for the people with hearing problems to communicate with other people. By recognizing the gestures, people with the hearing problems can communicate with one they wanted to without the need of mediators. The system which we wanted to build will capture the gestures made by the people specially the hand movements and based upon the symbols they framed a grammatically correct sentence and displayed in the form of captions. Due to the modern world and technology, everything needs to be upgraded so our project is a real-time live sign language detector, which is a step forward to communicate with the people who have hearing and speaking problems. We are trying to develop a live sign language detector which identifies the signs through the convolutional neural networks (CNN) and also converts it into meaningful sentences. It is basically a series of image processing techniques with human hand movement classification and recognition. This paper presents a recommendation model to identify and translate the hand gestures of sign language into meaningful sentences in English. The search is mostly based on the CNN models, TensorFlow and Keras libraries. No facial or body movements were used in our model and our project is dynamic in nature as it processes the gestures through the videos from a HD cam. Our study is conducting various gestures feed to the model and the prediction of the sign language. The project which we are aiming to build uses a machine learning technologies built using the RNN and CNN models. The paper is completely user-independent, and the classification machine learning algorithms are trained using very large datasets. The user whoever wants to communicate can just open our application and record the video of the person, and the corresponding text will be displayed on the screen which will have a particular meaning.

Keywords Convolutional neural networks · TensorFlow and Keras libraries · Gestures

C. Sri Sumanth (✉) · D. Nanda Surendhar · G. Pramod · P. Pranav Rao · T. Babu Karthik
V N R Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India
e-mail: sumanth.chekuri@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_28

1 Introduction

Sign Language is the majorly used medium to communicate with the people who are specially abled like dumb and deaf people. Sign language is specially organized language [1] which gives particular meaning to their corresponding particular symbol and the one who want to communicate uses these special symbols with their hands which can be seen and understood by those [2] who known the sign language, but the one who does not know the sign language has to use some or other interpreters to communicate with them.

Now a days with the enhancement of the technology [3], many advanced ways and techniques are developed so the specially abled persons can communicate comfortably. Consider a situation where any one wanted [4] to communicate with the deaf person and they do not know the sign language, it will be a very difficult situation for both of them if anyone of them do not know the sign language. The society thinks that the deaf and dumb persons are not important. We wanted to build a project which can help the deaf and dumb community. The software solution can be helpful in making them interact with the one they wanted.

Sign language is one crucial language which is used to communicate with specially abled people like dumb and deaf people. It uses facial expressions, body language and hand gestures to [5] communicate with specially abled people. The sign language recognition system can be classified into two types: they are static and dynamic [6] models. The static model is used for static images and predefined sentences, whereas the dynamic model is used for dynamic inputs like videos and constantly changing sentences.

In this model, we are using a dynamic system such that the model takes the live recording of a person using sign language and converts the sentences into a proper verbal language. ASL alphabets and ASL [7] numbers are used to train the model and test the live inputs against the trained model. This model benefits all the society because the society does not need to learn sign language to understand the dumb and deaf people when this model is used.

In this model, various machine learning algorithms are used and trained against large datasets to produce accurate results. ASL is one of the widely used languages in most of the deaf educational institutions. This [8] model helps around 18 million people who are deaf and speech-impaired people around the world to communicate with common people.

The sign language is also unique, each place has different meaning for a particular action. Many people customize the language as per their needs. There are many restaurants in the world [9] where each action indicates to a particular dish and the people working in that place are mostly deaf and dumb people and the applications like these will help them and will make their job simpler and also let them communicate with the other people [10] easier.

2 Literature Survey

In Ojha [11], they uses CNN model for gesture detection and trained using gestures dataset, various skin complexion and quick hand gestures yield inefficient or wrong output sometimes.

In Kollerq [12], they have used CTC (connectionist temporal classification) a single unified architecture that increases the performance significantly. This model is very complex and requires a very large dataset.

In Elmahgiubi [13], a smart glove is used to capture the gestures and they used a model to recognize those gestures and display them into readable text on external screen. This model requires an external apparatus. This gloves has to be custom-made for the user for accurate results.

In Amin [14], they have used ROUGE and BLEU 4 models, although these models are incomplete, they have showed superiority of the proposed models, which needs a further research.

In Shinde [15], the edges of the hand gestures is detected using the edge detection algorithm and then the model undergoes gray scaling to perfectly recognize the input image. It is a static model and requires to record gestures and feed it to the model. It does not give grammatically correct sentences.

In Papadogiorgaki [16], the model takes a sentence as input and gives a VRML 3D model as output. Torso movements are not supported and also some facial expressions like cheek wrinkles.

In Nagarajan [17], the CNN model is used to recognize the hand gestures of sign language through SVM multiclass and EOH—edge-oriented histogram. As the model is a static model, a small change in the adjustment in the image might vary the output sentence drastically.

In Lozanova [18], a rule-based system is used for automatic moderation of Bulgarian text to help the deaf people for proper education. The level of understanding is very low, and the complex set of alterations will have to be implemented to attain better overall results.

In Escudeiro [19], a translator is used for the conversion using Microsoft Kinect and 5DT sensor gloves. These external apparatus are used to improve portability and cost of implementation.

In Schmidt [20], a video of a person signing is translated into text of the spoken language, but the model uses Viseme recognition that is inaccurate and is of poor quality.

3 Methodology and Algorithm

The proposed system is given such that it detects the hand gestures through videos and coverts it into meaningful sentences. We are using CNN (convolutional neural networks) model for the gesture [21] recognition.

From the videos, the images are taken by the webcam dynamically feeding it to CNN model. We are using Keras, OpenCV Python for the architecture of our model.

The work flow of the system is given below.

3.1 Input Video

The images are randomly taken from the video and given as input to our CNN model. A hand histogram box will be appeared in which the hand gestures should be in those histogram square. Hand gestures are detected from the web camera [22], and these gestures are stored in our database for model training and preprocessing. The ASL alphabets and ASL numbers are used for gesture detection which will be used in training the model.

3.2 Preprocessing

Preprocessing of the model includes the gesture classification and storing them. We are using [23] CNN model so that we give the input by flipping them or reducing size of them. The preprocessing depends upon the size of the image, ratio of the image, etc. During the preprocessing stage, the images, i.e., the gestures are labelled [24] and given certain names for calculating the probability of the image.

3.3 Model Training

The training of our model is done using Keras. Very large datasets are used to train our algorithm. The model does not need to restore every time unless you are added or removed a gesture. For training the model, training dataset is used in which different gestures and their labels were given to obtain the accuracy of the model.

Convolutional layers and fully connected [25] layers are also depended upon the training dataset. During this stage, the hand gestures are converted into actual verbal conversation language. CNN model contains convolutional layer, dense layer, max pooling layer and fully connected layer [26] to produce accurate and precise conversion from image to meaningful sentence (Fig. 1).

3.4 Testing Gestures

The model is tested by real-time live gestures if the model is established, then we need to check for the lighting conditions. The output will be text format which is

Fig. 1 ASL alphabets and ASL numbers

meaningful displayed on the screen [27] when you remove the hand, and the entire report is documented and predicted including accuracy, confusion matrix and error rate.

4 Conclusion and Future Scope

Sign language recognition is very difficult to process as it a continuous change of gestures through videos that the system needs to understand and translate into a meaningful sentence. So, we decided to [28] solve this problem using variable machine learning concepts like CNN architecture to obtain a better performing recognition system that can be used to translate the sign language into proper verbal language.

This project uses ASL alphabets and ASL numbers as the database [29] training. The model captures continuous images from the videos and trains against the ASL numbers and alphabets to provide a proper meaningful sentence. This project is mostly useful for the persons who want to communicate with the dumb and deaf people even without knowing the sign language. This [30] model can be used in various industries like robotics, general day-to-day activities to understand dumb and deaf people. This model can be upgraded to give more efficient sentences in less time and can also be installed and used in various sectors.

References

1. Karthick P, Prathiba N, Rekha VB, Thanalaxmi S (2014) Transforming Indian sign language into text using leap motion. Int J Innov Res Sci Eng Technol 3(4):10906–10910
2. Bragg D, Kushalnagar R, Ladner R (2018) Designing an animated character system for american sign language. In: Proceedings of the 20th international ACM SIGACCESS conference on computers and accessibility. Association of Computer Machinery, United States
3. Grieve-Smith AB (2016) SignSynth: a sign language synthesis application using Web3D and Perl. In: Technische Fakultät international gesture workshop (2002). Springer, Heidelberg
4. Halawani S (2008) Arabic sign language translation system on mobile devices. Int J Comput Sci Netw Secur 8(1):251
5. Kahlon NK, Singh W (2021) Machine translation from text to sign language: a systematic review. In: Universal access in the information society, India
6. Duarte AC (2019) Cross-modal neural sign language translation. In: Proceedings of the 27th ACM international conference. Computer science Linguistics, United States
7. Borgna G, Convertino C, Marschark M, Morrison C, Rizzolo K (2010) Enhancing deaf students' learning from sign language and text: metacognition, modality, and the effectiveness of content scaffolding. J Deaf Stud Deaf Educ 16(1):79–100
8. Fenlon J, Schembri A, Rentelis R, Vinson D, Cormier K (2014) Using conversational data to determine lexical frequency in British sign language: the influence of text type. Lingua 143:187–202
9. Elliott R, Glauert JRW, Kennaway JR, Marshall I, Safar E (2007) Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. In: Universal access in the information society, UK
10. Ormel E, Giezen MR, Knoors H, Verhoeven L, Gutierrez-Sigut E (2002) Predictors of word and text reading fluency of deaf children in Bilingual deaf education programmes. Cogn Nat Bilingual Reading 7(1):51
11. Ojha A, Pandey A, Maurya S, Thakur A, Dayananda PI (2022) Sign language to text and speech translation in real time using convolutional neural network. Int J Eng Res Technol (IJERT), India
12. Camgöz NC, Koller O, Hadfield S, Bowden R (2020) Sign language transformers: joint end-to-end sign language recognition and translation. In: IEEE/CVF conference on computer vision and pattern recognition (CVPR), USA
13. Elmahgiubi M, Ennajar M, Drawil N, Elbuni MS (2015) Sign language translator and gesture recognition. In: Global summit on computer & information technology (GSCIT), Sousse, Tunisia, pp 1–6. https://doi.org/10.1109/GSCIT.2015.7353332
14. Amin M, Hefny H, Ammar M (2021) Sign language gloss translation using deep learning models. Int J Adv Comput Sci Appl 12(11)
15. Shinde A, Kagalkar R (2015) Sign language to text and vice versa recognition using computer vision in Marathi. In: International Journal of Computer Applications (0975–8887) National Conference on Advances in Computing (NCAC 2015)
16. Papadogiorgaki M, Grammalidis N, Tzovaras D, Strintzis MG (2005) Text-to-sign language synthesis tool. In: 13th European signal processing conference, Antalya, Turkey, pp 1–4
17. Nagarajan S, Subashini TS (2013) Static hand gesture recognition for sign language alphabets using edge oriented histogram and multi class SVM. Int J Comput Appl (0975–8887) 82(4), India

18. Lozanova S, Stoyanova I, Leseva S, Koeva S, Savtchev B (2013) Text modification for bulgarian sign language users. In: Proceedings of the second workshop on predicting and improving text readability for target reader populations, Sofia, Bulgaria. Association for Computational Linguistics, pp 39–48
19. Oliveira T, Escudeiro P, Escudeiro N, Rocha E, Barbosa FM (2019) Automatic sign language translation to improve communication. In: Published in IEEE global engineering education conference (EDUCON), Dubai, United Arab Emirates, pp 937–942. https://doi.org/10.1109/EDUCON.2019.8725244
20. Schmidt C, Koller O, Ney H, Hoyoux T, Piater J (2013) Using viseme recognition to improve a sign language translation system. In: Proceedings of the 10th international workshop on spoken language translation: Papers, Heidelberg, Germany
21. Chen Y, Wei F, Sun X, Wu Z, Lin S (2022) A simple multi-modality transfer learning baseline for sign language translation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)
22. Marschark M, Leigh G, Sapere P, Burnham D, Convertino C, Stinson M, Knoors H, Vervloed MPJ, Noble W (2006) Benefits of sign language interpreting and text alternatives for deaf students' classroom learning. J Deaf Stud Deaf Educ 11(4):421–437
23. Boulares M, Jemni M (2012) Mobile: sign language translation system for deaf community. In: International cross-disciplinary conference. Association for Computing Machinery
24. Parton BS (2005) Sign language recognition and translation: a multidisciplined approach from the field of artificial intelligence. J Deaf Stud Deaf Educ 11(1):94–101
25. Krausneker V, Becker C, Audeoud M (2020) Bilingual school education with spoken and signed languages in Europe. Int J Bilingual Educ Biligualism 5(6):1794–1810
26. Austin AC, Schuler KD, Furlong S, Newport EL (2021) Learning a language from inconsistent input: regularization in child and adult learners. Lang Learn Dev 18(3):249–277
27. Hernandez-Rebollar JL, Kyriakopoulos N, Lindeman RW (2004) A new instrumented approach for translating American Sign Language into sound and text, IEEE. Seoul, Korea (South)
28. Sharma P, Tulsian D, Verma C, Sharma P, Nancy N (2022) Translating speech to indian sign language using natural language processing. Trends Data Sci Knowl Discov 14(9):253
29. Zhang Z, Wen F, Sun Z, Guo X, He T, Lee C (2022) Artificial intelligence-enabled sensing technologies in the 5G/internet of things era: from virtual reality/augmented reality to the digital twin. Adv Intell Syst
30. Kunjumon J, Megalingam RK (2019) Hand gesture recognition system for translating Indian sign language into text and speech. In: International conference on smart systems and inventive technology (ICSSIT), India

Utilizing Deep Natural Language Processing to Detect Plagiarism

K. Praveen Kumar, D. Jaya Kumari, and P. Uma Sankar

Abstract Plagiarism is now pervasive in a variety of spheres of life, including academia and research. The development of plagiarism strategies used by plagiarists makes it difficult for existing approaches to accurately detect plagiarism. Plagiarism is checked using a variety of aspects, including syntactic, lexical, semantic, and structural features. This study examines novel and contemporary plagiarism detection tasks, particularly text-based and monolingual plagiarism detection. We suggested a four-stage innovative approach for detecting plagiarism. The natural language processing (NLP) methodology is used in this framework as opposed to the more conventional string-matching methods. By combining two metrics—skip gram and dice coefficient—on the basis of a corpus-based approach, this system investigates text similarity. Using the deep and shallow NLP approach, the text's deeper meaning is investigated. Our findings indicate that deep NLP is swiftly recognizing heavy revision. Shallow NLP efficiently prepares text for future processing. The findings of Word2vec are comparable to those of straightforward deep NLP techniques, however Word2vec also emphasizes documents that other methods might miss. Deep NLP also records changes in synonyms and phrases.

Keywords Natural language processing · Skip gram · Dice coefficient · Plagiarism · Word2Vec · Shallow NLP · Deep NLP

K. Praveen Kumar (✉) · D. Jaya Kumari · P. Uma Sankar
Department of Computer Science and Engineering, Sri Vasavi Engineering College(A), Pedatadepalli, Andra Pradesh, India
e-mail: praveenkumar.cse@srivasaviengg.ac.in

D. Jaya Kumari
e-mail: hod_cse@srivasaviengg.ac.in

P. Uma Sankar
e-mail: umasankar.cse@srivasaviengg.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_29

1 Introduction

The act of using another person's original words and ideas as one's own ideas and viewpoints is known as plagiarism. Since the advent of the Internet [2], it has been simple to swiftly access a large amount of data, making the possibility of copying, pasting, or otherwise utilizing the words or ideas of others more likely. Without formal references, it is an unlawful act to force anyone to labor as their property. Individuals that plagiarize are thieving, stupid, and lazy in both their academic [3] and professional lives. Plagiarism is associated with issues that are difficult and complex. Plagiarism is on the rise as well as Internet usage. In any case, it is challenging to implement comparative techniques in regular writing, thus it is difficult to use this strategy to make a text original. For this reason, both free and paid tools for detecting plagiarized work have been developed and are in use. Such systems still struggle to reliably identify cases of phrase modifications or sentence rewriting. Our research [1] aims to develop a method that detects plagiarism in monolingual scenarios while overcoming all current difficulties. We provide a framework that combines the word embedding deep NLP technology, the skip gram [7], and the dice metric to examine complex plagiarism more precisely. Four stages make up this framework: preprocessing, similarity comparison, filtering, and document classification.

The quick advancement of science and the rise in crime is just two effects of increased internet use on social lives. One of these is plagiarism, which is when a piece of a material that has been protected by copyright is used in its entirety without properly citing the author(s). It is also referred to as a type of intellectual property theft that doesn't accurately credit the source document's creator but yet closely matches the idea of the original work by duplicating it intrinsically or extrinsically. 67.4% of 6096 undergraduate students [8] from 31 universities who participated in the study were determined to have plagiarized. More than 6000 high school and college students participated in a comparable survey on numerous different campuses, and the findings revealed that 76% of them had plagiarized [2]. There are various approaches for detecting plagiarism, but not all types of plagiarism may be found. Intrinsic and extrinsic plagiarism are the two categories into which plagiarism is separated. Near copies, disguised, translated, and idea are the four types of intrinsic plagiarism that need to be taken into account while detecting plagiarism. Near copies, a type of plagiarism, involve reproducing (exactly the same as) the source without properly attributing it.

2 Literature Survey

In this section, we will mainly discuss about the background work that is carried out in order to prove the performance of our proposed method. Literature survey is the most important step in software development process. For any software or application development, this step plays a very crucial role by determining the several factors like time, money, effort, lines of code, and company strength.

Parker and Hamblen et al., long ago [14], discussed about the plagiarism detection using several algorithms. However, given the volume of data and the sophistication of the tactics employed to evade detection mechanisms, it continues to be a real issue. A substantial amount of work has been published starting from the five elements of plagiarism established in Parker and Hamblen [14] and moving on to the real approach of deep learning described in Bakhteev et al., although there are still unresolved issues. Choi and Ko. According to Lancaster et al., techniques can be divided into categories based on the sort of detection methodology used, the system's accessibility, the volume of documents the metrics can analyze, and the complexity of the metrics. It has been highlighted that pair-wise measurements with low structural complexity have been used the most frequently, as opposed to more precise multi-dimensional metrics. This is brought on by the compromise between accuracy and processing power. Even with the use of powerful computers, performing detection tasks would frequently take a long time and a lot of effort due to the complexity of the measurements. For people with personal computers, this is not the best situation.

Although Marsh and Williams et al. suggested Turnitin, as the most successful commercial detection tool, endorsed by JISC1 as a teaching tool for plagiarism prevention and detection, user reviews of the tool have not been positive at all because Turnitin is unable to handle paraphrased texts well [3]. Ahmed Hamza Osman and Omar M. Barukab et al. in [4] proposed a semantic plagiarism detection method in light of support vector machines (SVM) and semantic role labeling (SRL). Comparing the time-effectiveness of graph-based with SRL-SVM, semantic similarity detection technique,fuzzy semantic string similarity, and LCS in terms of time effectiveness. The time efficiency of the suggested method was also established; it belongs to the O(n2) Class [9]. NLP involves the processing of human languages by using machines and this is included in many fields such as computer-assisted language learning and extraction of biomedical information has already experienced benefits from using NLP. However, it remains an under-explored area for plagiarism detection.

3 Existing Methodology

Currently, there is no technique to check for plagiarism by comparing one file with other file formats such as PDF and jpeg. Whatever system is in place for detecting plagiarism currently only works for documents or text files, or easily take, we may compare similarly formatted files. There is no mechanism that can detect plagiarism for photos; all the systems that are now in use attempt to detect plagiarism for text data and check it either from the Google repository or from other data sources to estimate the proportion of plagiarism.

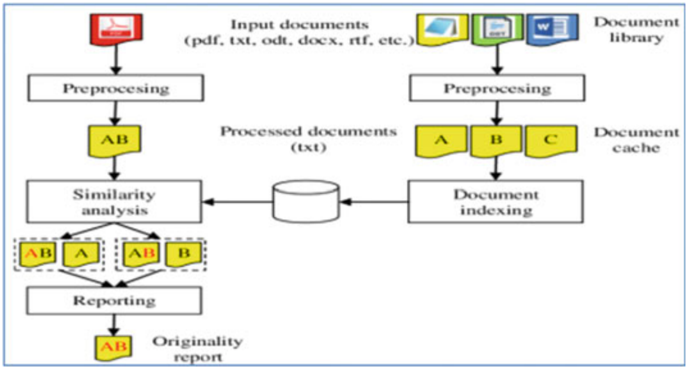


Fig. 1 Primitive text plagiarism detection

3.1 Limitation of Existing System

1. A longer wait time before plagiarism can be done.
2. Every approach now in use looks for instances of text data plagiarism.
3. There is currently no machine learning method that can quickly detect plagiarism in PDFs or photos.
4. If a file's input contains a lot of complex information, plagiarism detection softwares currently on the market do not produce accurate findings.
5. If the input document contains copied images, the text cannot be extracted from the images to determine whether or not the text is original.

Figure 1 clearly illustrates the primitive architecture. The major focus of text plagiarism detection is on text files. Generally speaking, we strive to accept input as either. Word, PDF, and so forth. After the software receives these files as input, it will attempt to perform a similarity analysis [10] on the documents based on the document library. Once the similarity is discovered, we will be able to obtain a report for that content based on the amount of words that match the document library. The user will see this reflected as a percentage at the conclusion of the program.

4 Proposed Methodology

In this stage, we are going to explain the proposed system and its model by using some model diagram and now we can clearly identify the step by step procedure of our proposed system. In the proposed system, one reference text file [4] is used as a starting point, and then plagiarism is checked using various file types, such as txt, PDF, or jpg. To convert PDF information into text, we use the pypdf2 module, and to convert image information into text, we use easy OCR. We also pre-process text information in four stages as listed below, and compare the document similarity score with the reference file [10].

(1) **Preprocessing is the first stage**: For later steps, the source text and any potentially suspicious documents are ready. The messages are subjected to text preprocessing and superficial natural language processing (NLP) approaches [5].
(2) **Dice and Skip gram metrics are the Second Stage**: Dice and skip gram metrics are used to compare similarities. Using skip gram and the dice coefficient metric, source and suspicious documents is compared pair wise. The following stage will then continue processing these computed scores.
(3) **Deep NLP Filtering is the Third Stage**: Deep NLP filtering candidate pairs are then generated once the deep NLP approach and word2vec process have been reviewed. By setting a threshold at each similarity score, the analysis of similarity scores creates the candidate pair. A pair is designated as a candidate pair if it matches the chosen threshold. In this method, by examining deep NLP first, those documents with complicated plagiarism that were disregarded owing to low similarity values are also studied. Afterward, using the threshold point as a guide, we can select the candidate documents.
(4) **Classification is the fourth Stage**: The selection of plagiarized or non-plagiarized material is indicated here after candidates and similarity scores have been chosen from the previous step. The WEKA tool can be used to categorize documents.

On a sizable amount of data from the publicly accessible dataset, the suggested model is trained and tested. The model can accept input in the form of text files, pictures, word processing files, or PDF files. It can then attempt to apply deep NLP approach to the dataset and, as a final step, check classification. Finally, we attempt to categorize whether or not the document involves plagiarism (Fig. 2).

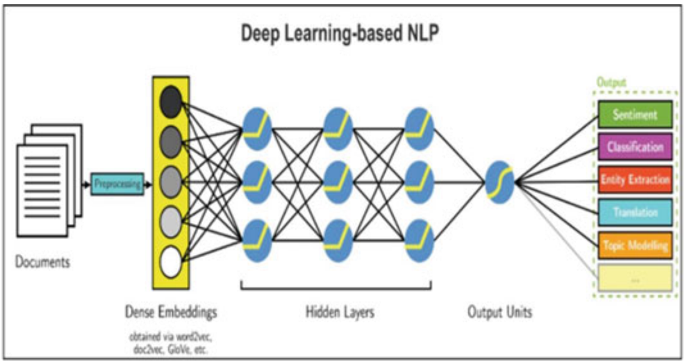


Fig. 2 Deep learning NLP for plagiarism detection

5 Proposed Deep Natural Language Processing Model

The proposed methodology attempts to use deep NLP, a subset of AI, to determine whether or not there is plagiarism in the supplied material. The rudimentary plagiarism detection tools are typically unable to process the photos and determine whether they match in any of the document resources.

Skip Gram

In order to make computers more like people, the artificial intelligence area of natural language processing (NLP) aims to process human language as text or speech. Humans possess a significant amount of information that is poorly arranged. Therefore, it's challenging for any machine [6] to interpret raw text. We must convert this data into a vector format so that our computers can quickly process it in order to make a machine learn from the raw text. Word representation is the process of converting plain text into a vector format [11].

When a word's vectors are close to one another, it indicates that the words are connected to one another because the word representation depicts words in vector space. In the Fig. 3, the terms that are typically associated with women are grouped on the left, while the words that are typically associated with men are grouped on the right. Therefore, if we input the term "earrings," the computer will logically associate it with the female gender. Due to the size of any language's vocabulary and the difficulty of human labeling, unsupervised learning techniques are needed so that any word can understand its context on its own. One of the unsupervised learning methods employed is called the skip gram. To determine the context word for a given target word, skip gram is employed. It is the CBOW algorithm inverted. Here, the target word is entered and the surrounding words are produced. This problem is challenging because there are multiple context words that must be predicted [12].

Fig. 3 Word embedding in the NLP for plagiarism detection

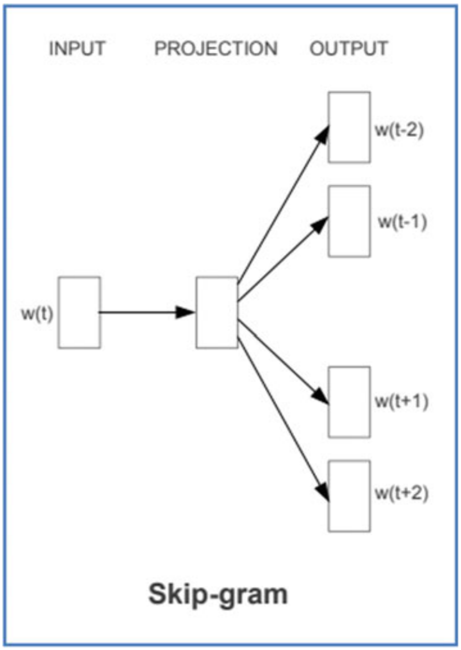


Fig. 4 Skip gram architecture

As we can see Fig. 4, $w(t)$ is the given input or target word. One hidden layer computes the dot product of the input vector, w, and the weight matrix (t). In the hidden layer, no activation function is employed. The output layer is now given the result of the dot product at the concealed layer [13]. The output layer computes the dot product of the weight matrix of the output layer with the output layer's output vector. The likelihood of words existing in the context of $w(t)$ at a given context location is then calculated using the softmax activation function.

From the above Fig. 5, we can clearly identify the input will be undergoing some internal processing in order to get the desired output. In this proposed work, we used skip gram model for calculating the probability and loss function on any input. Here, we try to use this skip gram on plagiarism detection on text and image formats, the following are the equations for calculating probability and loss function.

Probability Function

$$p\left(w_{c,j} = w_{O,c} \middle| w_I\right) = \frac{\exp u_{c,j}}{\sum_{j=1}^{V} \exp u_j}$$

The single input word is $w(I)$, the jth predicted word for the cth context position is $w(c,j)$, the actual word for the cth context position is $w(O,c)$, and the jth value in the U vector when predicting the word for the cth context location is $w(c,j)$.

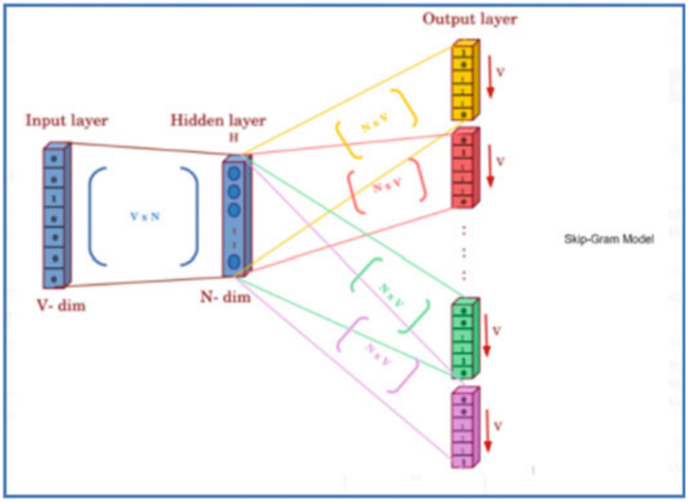


Fig. 5 Skip gram internal flow

Loss Function

$$\mathcal{L} = -\log \mathbb{P}\left(w_{c,1}, w_{c,2}, \ldots w_{c,C} \middle| w_o\right) = -\log \prod_{c=1}^{C} \mathbb{P}\left(w_{c,i} \middle| w_o\right)$$

$$= -\log \prod_{c=1}^{C} \frac{\exp\left(u_{c,j^*}\right)}{\sum_{j=1}^{V} \exp\left(u_{c,j}\right)} = -\sum_{c=1}^{C} u_{c,j^*} + \sum_{c=1}^{C} \log \sum_{j=1}^{V} \exp\left(u_{c,j}\right)$$

We can express the loss function $\mathcal{L}$ because we wish to increase the likelihood of correctly guessing $w(c,j)$ on the cth context location.

Dice Coefficient

The Srensen-Dice coefficient, a statistic created in the 1940s to assess the similarity between two samples, is where dice loss is named. In 2016, Milletari et al. presented it to the computer vision community for 3D medical image segmentation.

In our current application, we try to find out the plagiarism from images and hence the images need to be segmented first and then from those images we are going to collect the information. During this process, some of the information may not be extracted properly so we may get loss, hence this loss can be calculated using the dice coefficient.

The dice coefficient is 2 *the area of overlap divided by the total number of pixels in both images. Here, when we try to extract the content from images, some portion of content will get lost such as edges which is shown in Fig. 6, background and so on. Hence, we need to use the dice coefficient to calculate the loss function.

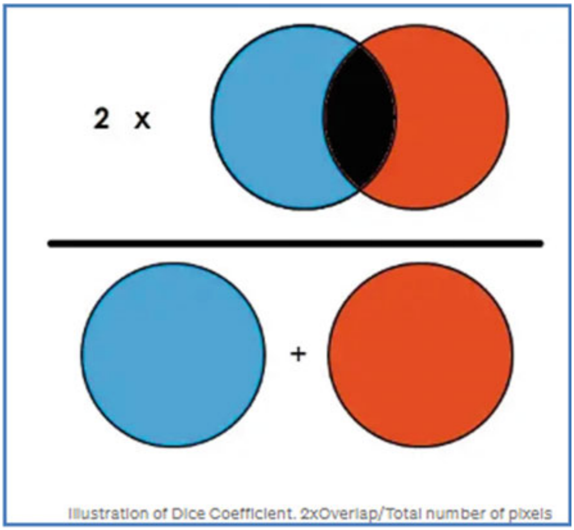


Fig. 6 Illustration of dice coefficient

6 Result and Discussion

In this proposed work, we try to utilize the python as programming language to construct the application and Google Collaboratory is used as work platform to execute the application. We collected the dataset from

https://data.mendeley.com/public-files/datasets/gz3hztwm5p/files/e6762290-63ea-4682-a4a3-291dd479bc8d/file_downloaded

The above link is public URL in which lot of text, images, and PDF contents are present which contains normal and plagiarized content inside the files. Hence, we try to use this dataset for finding the plagiarism using deep NLP classifications.

In the above Fig. 7, we can clearly see deep NLP model is applied on the input dataset and then try to process the model and now we try to test the model on training data. Hence, we divide the input dataset into test and train (80% train and 20% test) (Fig. 8).

In the above Fig. 9, we can clearly see we gave some text documents as input for our application which contains text as well as images and then the files which contain plagiarism is identified in the performance evaluation window which is seen in Fig. 10. The above bar graph makes it evident that doc3 has a higher percentage of plagiarized material than docs 1 and 2. As a result, doc3 is regarded as a plagiarized work. This doc3 can have internal graphics or symbols with unique notations. However, our suggested application can predict the percentage of plagiarism with accuracy.

Apply Deep NLP Model

```
'SADC Energy Ministers',
'Executive Committee',
'Management Committee',
'Planning',
'Operating',
'Environmental',
'Sub-committee',
'Sub-committeel',
'Sub-committee',
'Co-ordination',
'Centre']
```

```
len(sorted(os.listdir('Source_figures/')))
```

```
298
```

```
!mkdir Source_text_files
!cp Source_figures/*out.txt Source_text_files/ # copying the text files into separate folder
!mkdir Source_images
!cp Source_figures/*.jpg Source_images/ # copying the image files into separate folder
! rm Source_text_files/Source_58_00_out.txt
!rm Source_images/Source_58_00.jpg
```

```
len(sorted(os.listdir('Source_text_files/')))
```

Fig. 7 Deep NLP model

Test & Train Validation

Testing the documents

```
from bs4 import BeautifulSoup
from nltk.tokenize.treebank import TreebankWordDetokenizer
import re
import nltk
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('punkt')
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()
def clean_input(sent):
    # print(sent)
    #remove html content
    review_text = BeautifulSoup(str(sent)).get_text()

    #remove non-alphanumeric characters
    review_text = re.sub("[^a-zA-Z0-9]+"," ", review_text)
    review_text=review_text.lower()
    words = word_tokenize(review_text)
    stops = set(stopwords.words("english"))
    #
    # 5. Remove stop words
    meaningful_words = [w for w in words if not w in stops]

    #lemmatize each word to its lemma
    lemma_words = [lemmatizer.lemmatize(i) for i in meaningful_words]
    sentence=TreebankWordDetokenizer().detokenize(lemma_words)
```

Fig. 8 Test and train

Test Input File

Fig. 9 Test input file

Performance Evaluation

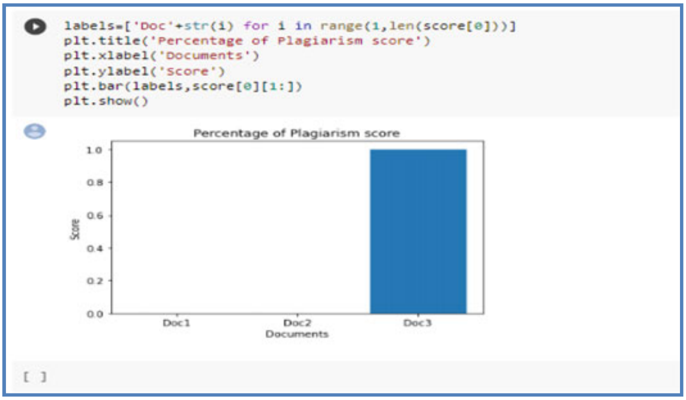


Fig. 10 Detect plagiarism

7 Conclusion

This study presents the findings of a novel algorithm to detect plagiarism more precisely, one that can handle both intrinsic and extrinsic plagiarism. This demonstrates that the algorithm used to identify plagiarism below may yield more accurate results with a number of changes: (1) The use of semantic parsing increases the accuracy of detecting plagiarism, especially in paraphrased plagiarism. (2) The use of syntactic parsing, in this case POS-tagger, in identifying plagiarism. The more appropriate words to obtain will result from utilizing this method in conjunction with semantic analysis to locate related terms, and (3) arranging the words in a phrase before performing the matching will result in an efficient use of time.

References

1. Kashkur M, Parshutin S (2010) Research into plagiarism cases and plagiarism detection methods. Riga Tech Univ Sci J 44(1):138–143
2. Scanlon PM, Neumann DR (2002) Internet plagiarism among college students. J Coll Stud Dev 43(3):374–385
3. Meuschke N, Gipp B (2013) State of the art in the detecting academic plagiarism. Int J Educ Integrity 9(1):50–71
4. Understanding plagiarism linguistic patterns, textual features, and detection methods. IEEE 42(2):133–149
5. Mihalcea R, Liu H, Lieberman H (2006) Proceedings of the international conference on computational linguistic and intelligent systems. In: (NLP) Natural language processing for (NLP) natural language programming, Text Processing (CICLing), Mexico City (Mexico), 19–25 Feb 2006, pp 319–330
6. Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P (2011) Natural language processing (almost) from scratch. J Mach Learn Res 12:2493–2537
7. Stamatatos E (2009) Intrinsic plagiarism detection using character n-gram profiles. In: Proceedings of the Spanish society for natural language processing (SEPLN) international conference, San Sebastian (Spain), 8–10 Sept 2009, pp 38–46
8. Automatic student plagiarism detection: future perspectives. J Educ Comput Res 43(4):507–527 (2010)
9. Botana G, Leon J, Olmos R, Escudero I (2010) Latent semantic analysis parameters for essay evaluation using small-scale corpora. J Quant Linguist 17(1):1–29
10. Micol D, Munoz R, Ferrandez O (2011) Investigating advanced techniques for document content similarity applied to external plagiarism analysis. In: Recent advances in natural language processing (RANLP) conference proceedings, Hissar (Bulgaria), 12–14 Sept 2011, pp 240–246
11. Using natural language processing for automatic detection of plagiarism. In: Proceedings of the international plagiarism conference (IPC), Northumbria University (Newcastle), 21–23 June, 2010
12. Kucecka T (2011) Plagiarism detection in obfuscated documents using an N-gram technique. ACM 3(2):67–71
13. Bose R (2004) Natural language processing: current state and future directions. Int J Comput 12(1):1–11
14. Parker A, Hamblen J (1989) Computer algorithms for plagiarism detection. IEEE Trans Edu 32:94–99. http://dx.doi.org/10.1109/13.28038

Pocket-Pi a Low-End Learning and Deploying Tool for Intelligent Systems and IOT

N. N. S. S. S. Adithya, Cheryl D. Souza, and Gurram Shravya

Abstract As far as the artificial intelligence and machine learning applications are consider, they are two most important challenges deploying such intelligent systems into real-world environment. One being the compactness and secondly the integration with IOT-based solutions. In this paper, we propose a simple easy-to-use plug-and-play device which supports both learning and deploying small- to mid-scale projects based on AI/ML and IOT across various environments. The proposed solution will be removing the barrier to learn and deploy programming applications. A single-board computer (SBC) is equipped with all the infrastructure such as battery for power, heat sink for thermal conductivity NVMe M.2 solid-state drive for storage and is supported by all necessary software packages like OpenCV, TensorFlow, NumPy, Matplotlib, pandas and SciPy, which are natively compiled for the ARM BCM2711 processor in a customized operating system.

Keywords IOT · SBC · Artificial intelligence · ARM · Solid-state drive · Machine learning

1 Problem Statement

The major hurdle for any people who wants to learn programming and deploying application on AI/ML in any language will be the installation and configuration of the programming environment which he desired to learn and use. The major PC and laptops that they use may not have the specification required for that environment as most of the people are still using an older version of hardware, because of new

N. N. S. S. S. Adithya (✉) · C. D. Souza
Department of Computer Science Engineering, CVR College of Engineering, Ibrahimpatnam, India
e-mail: adithya365.ii@gmail.com

G. Shravya
Department of AI and ML, Malla Reddy Institute of Technology and Science, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_30

machines becoming quite expensive, and due to the recent advancement in technologies, there is a rapid development in the supportive hardware such as GPU and CPU which makes them very efficient. Current chip shortage and scale-up prices make them far away from our reach; thus, an affordable and upgradable device is needed with supporting majority of programming platforms and libraries like TensorFlow, Keras and SciPy. This device must also be able to perform regular computer works such as spreadsheets, documents, PDFs, power point presentations, media playback and Internet surfing.

The major challenge in our objective is to transform the technology which is sufficient for programming to remotely run-on other devices without any dependencies. SBC (single-board computers) are very good to start learning programming and making projects, but they are not quite user-friendly and powerful enough to replace laptops and computers. This device is very efficient for developing and deploying small-scale applications like IOT gateways, robotics, automations, artificial intelligence/machine learning solutions, and monitoring and reporting sensory information.

As they are providing a limited functionality, they need supporting equipment such as like power, storage, thermal management, which need to be purchased separately. For the software, the open source and communities are very helpful. The proposed solution packs, which includes all supporting hardware into a single unit, take care of all software needs and provide a code base which acts like a walkthrough for all the concepts and applications that you can develop through the device.

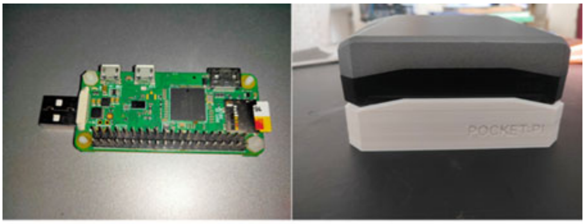
2 Related Work

PiTop is a London-based company which makes Raspberry Pi-based laptops, computers and stem-based products. They offer a complete solution to the infrastructure problem that exists in the Pi eco-system thought they offer a versatile solution in terms of connectivity. The major challenge is the remote connectivity type which Pi-Top uses VNC (virtual network connection) that will stream the entire Pi environment to main machine. Keeping with the desire to make STEAM learning engaging at all levels, the Pi-Top team has been working with educators from around the world, to better understand their needs and how we can help serve them. Beginning at the K-12 student level, all the way through to advanced computer science concepts, we have developed courses, challenges and teaching material to assist learning, which can be paired with all our products or, in some instances, can be followed on their own. They have wide range of products like robotics, electronics and other kits. As they are into more kind of educational space, we intended our focus in application development and deployment centric.

3 Proposed Device

The solution for this major problem, which making students to benefit from learning of programming, will be a Raspberry Pi-based plug-and-play device to any computer or laptop with all programming environments such as **C**, **C++**, **Java**, **Python**, and editor plugin for one of the popular programming editor **VS-Code**, which remotely logins and execute our programs in the plug-and-play device. The device can be operated in two ways either as stand-alone device by providing its own monitor, keyboard, mouse and power supply or by using an Ethernet cable which connects Pi to a computer and allows to have a remote access to the entire system, the built-in battery which can last up to 2–3 h of maximum usage. A solid-state drive is used for data storage and program space to user. The aim is to bring a change for every student who want to learn, experiment and make codes, which let them to become professional in programming, and to deploy simple-scale applications with minimalist interface such as billing services, cloud and NAS (network-attached storage) applications.

3.1 Hardware Unit



Elite version Ultimo Version

As per different requirements and consideration for different purposes, two models have been developed.

3.1.1 Elite Version

The elite version is very small and low-powered device, which is used to access the core Linux terminal, its main functionality is to provide a minimal environment, thus not requiring any additional power requirements. This can be used as USB plug in device and desktop or laptop. This version uses Pi-zero W, which is having 900 MHz processor and 512 MB of RAM, and this only supports remote SSH for shell-based scripts to execute and basic console-based programming. It is quite used fully to deploy familiarizing about Linux environment, deploying key loggers and rubber ducky programs, remotely monitoring network logging devices. A 40 GPIO (general purpose input output) pins for IOT and embedded applications are quite essential when low-powered device is needed, and for communication it includes Bluetooth and Wi-Fi and CSI (camera serial interface) connector for a Pi-cam, which can be used for simple image processing tasks like face recognition and object detection as it can run TensorFlow lite.

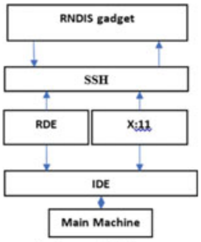
3.1.2 Ultimo Version

The ultimo version is a high-end device with latest Pi 4 Model B with 1.18 GHZ processor, 4 GB RAM and 64-bit OS. A passive heat sink made with brushed aluminum, 128 GB NVME solid-state drive for program space and robust body which is fully enclosed. This packs device for heavy performance-oriented tasks such as machine learning and computer vision-based operations. The device is equipped with 6600 mah which will be used when the device used as an extension to an existing system. The supportive hardware of this system is easy to upgrade by changing entire SBC which can improve the CPU, RAM and GPU at once an 8 GB model Pi will give you more RAM, and overclocking can increase the processor clock, as a result it can also increase thermal throttling which is handled by the massive heat sink for better thermal dissipation from the processor and graphics processing unit also at once. The GPIO pins has been labeled, and a magnetic top is used to cover to prevent dust clogging into the device. A power supply is mandatory for this version as it used heavy performance, and it provides a great price per performance rate.

3.1.3 Software Bundle

As far as software programming is considered, OpenCV 4.5.4 is compiled from source to run on the device by which user can easily run computer vision-based applications. TensorFlow 2.2.0 has been also build in same way along with PyQt5 for GUI development and SciPy form more complex ML development. All major concepts such as OOPs-based ML and AI applications have pre-loaded for simple understanding.

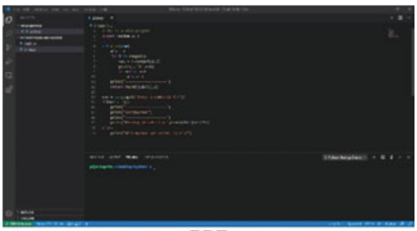
3.1.4 RDE Environment



System architecture

1. **RNDIS gadget**: Remote Network Driver Interface Specifications is a Microsoft proprietary protocol. Devices using USB connection for downloading and debugging of OS image from Platform Builder use this driver to emulate a network connection.
2. **SSH**: SSH is a protocol for securely exchanging data between two computers over an untrusted network. SSH protects the privacy and integrity of the transferred identities, data and files. It runs in most computers and in practically every server. It ships standard on UNIX, Linux, and macOS machines and it is used in over 90% of all data centers in the world.
3. **RDE: Remote Development** Environment allows you to use a container, remote machine or the Windows Subsystem for Linux (WSL) as a full-featured development environment.

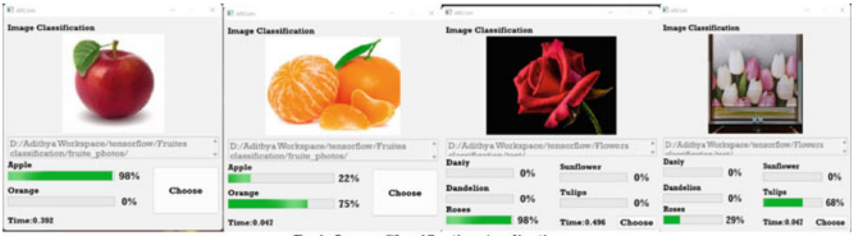
   X: 11: X: 11 is an architecture-independent system for remote graphical user interfaces.
4. **X: 11**: X: 11 is an architecture-independent system for remote graphical user interfaces and input device capabilities. Each person using a networked terminal can interact with the display with any type of user input device.



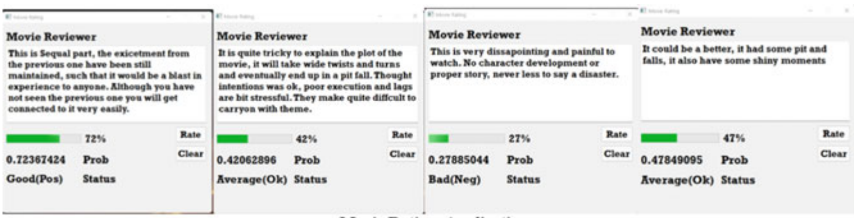
RDE

4 Results

Tensor flow Object Detection



Basic Image Classification Application



Movie Rating Application

5 Conclusion and Future Work

As our focus has primarily been to develop a robust hardware, which will be leverage the software packages and platforms we have provided. This system eliminates the mess up created by the supportive equipment and bring the simplicity with the user friendliness. A wide variety of applications like network-attached storage, Plex server, arcade emulation, media servers, IOT Gateways, robotics, automation system, conveyer belt automation and AI and ML applications and many more are deploying platform. This can be integrated to hardware using GPIO pins. Our hardware platform can be accessed from remote thus providing it OTA (over the air) updates.

Though the proposed system is quite effective in terms of scaling production and manufacturing, this proof of concept fails. We need to move our device with more optimized hardware platform which is **Compute Module 4** with customized **PCB** which reduces our power consumption, thermally efficient and reduces the manufacturing cost of the device. These advancements can take our prototype to production-ready device which can make integration into smart home devices and enterprise-level applications.

Implementation of Massive MIMO Technology with Artificial Intelligence Assisted Deep Learning Convolutional Neural Network (DLCNN)-Based Channel Estimation

Ch. Navitha and P. Anuradha

Abstract Fifth generation wireless communication requires the higher data rates to meet the requirements of real-world applications. However, the conventional multiple-input-multiple output (MIMO) technology unable to meet these requirements due to low performance channel estimation methods. Therefore, this article is focused on implementation of Massive MIMO technology with artificial intelligence assisted deep learning convolutional neural network (DLCNN)-based channel estimation. In Massive MIMO environment, channel is affecting by various types of uncertainties like noise, fading effects, and multipath propagations, which resulting reduced channel estimation performance in receiver side. Thus, the DLCNN model is trained with the different types of channel conditions and estimated the perfect channel response matrix. The simulations performed using MatllabR2022a shows that the proposed DLCNN channel estimation resulted in superior spectrum efficiency, energy efficiency, and base station density performance as compared to traditional channel estimators.

Keywords MIMO · Fifth generation · Massive MIMO · DLCNN model

1 Introduction

Wireless communication technology has witnessed significant improvements especially in the past decade, as new services and applications were launched and continuously being launched at increasing pace [1]. To effectively utilize these new services and applications, the seamless wireless connectivity has turned into an essential piece of our life [2]. The community has already observed the revolution in streaming media and because of its services like cable TV, music/movies on demand, and many

Ch. Navitha (✉) · P. Anuradha
Department of ECE, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: navitha_ece@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_31

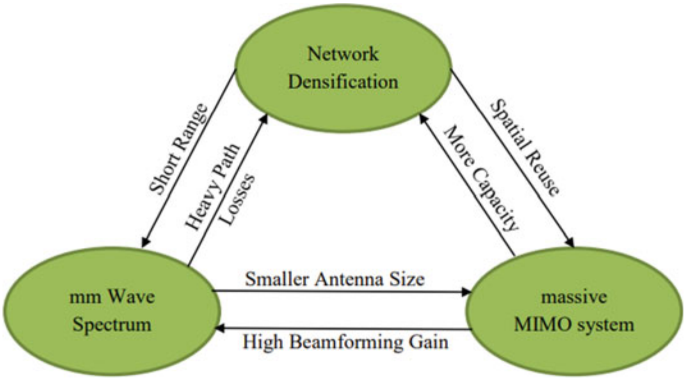


Fig. 1 Relation between three design aspects for upcoming wireless communication systems

more applications [3]. In addition, the new services/applications of smart cities, smart healthcare, smart wearable's, connected vehicles, intelligent transportations, connected machines, augmented reality, industrial IOT will need the uninterrupted wireless services almost on every corner of the world [4]. Consequently, the capacity of the system has to increase at exponential pace and the next generation 5G communication system is considered as the key technology enabler for enhancing the system capacity with uninterrupted connectivity and high service quality [5]. Figure 1 shows the relationship between Massive MIMO systems, mm-wave spectrum, and network densification. Here, the Massive MIMO shows that the prominent solution for fifth generation communications.

The conventional MIMO systems are failed to meet the maximum data rates in the fourth-generation networks [6] and the Massive MIMO hybrid channel estimation architecture for millimeter wave communication is key enabling technology for the upcoming wireless communication systems [7]. The hybrid channel estimation architecture for Massive MIMO systems is capable of providing the optimal spectral efficiency with relatively fewer number of RF chains which will reduce the overall power consumption as well as the cost of the communication equipment. The improvements in Spectral Efficiency are due to the use of spatial multiplexing techniques [8]. In spatial multiplexing, multiple symbols are transmitted over the same frequency and time slot. In addition, the transmitted energy is focused in the desired direction to enhance the SNR ratio and the process is called channel estimation [9]. Fully-digital channel estimation techniques are associated with high power consumption and high cost [10] as each antenna element of the array is connected with independent radio frequency chain which controls the amplitude also called as weight and phase of the signal fed to each antenna element. Due to these facts, the implementations of same for Massive MIMO is complex and expensive [11] although it provides the highest level of flexibility among all channel estimation techniques. In comparison to digital channel estimation, analog channel estimation requires less hardware, hence less power and is quite cost effective, as it requires only one RF chain per antenna array with total number of phase shifters equal to number of antenna elements [12]. Therefore, this work emphasizes on the major contributions in the field of the next generation wireless communication networks using Massive MIMO systems with DLCNN-based channel estimation.

Rest of the article is organized as follows: Sect. 2 deals with the literature review, Sect. 3 deals with the proposed system model, and Sect. 4 deals with results with discussions, and Sect. 5 deals with conclusions with future possibilities.

2 Literature Review

In [13] authors provided the extended review of use of millimeter waves for 5th generation wireless networks and the focus of paper is on propagation models. The authors initially preceded the key concepts of upcoming 5G wireless networks and available licensed as well as unlicensed millimeter wave band are described for various future applications. Further the propagation challenges [14] associated with each millimeter wave band are described in detail and associated antenna technologies are also elaborated. The propagation parameters and channel models proposed by prominent international groups are presented.

In [15] authors focused on the key point that the use of multi-beam antenna arrays in millimeter wave spectrum is the key technology to address the requirements of next generation cellular communication. The authors provided an extensive review of various multi-beam antenna technologies that can provide enhanced spectral efficiency and energy efficiency. In [16], authors covered both active as well as passive antenna technologies including their design, implementation, operating principle, and applications. In the end, preferred technologies for certain applications and associated challenges are highlighted.

In [17] authors proposed a hybrid channel estimation solution for a base station with high-dimensional antenna array. For the design, authors considered a single cell-MIMO system with perfectly known Channel State Information (CSI). The authors jointly designed the digital as well as analog precoder instead of traditional two-stage design approach. In [18], authors proposed solution for this CSI problems, which achieved impressive results for base station having large number of antennas as well as for lower-dimensional MIMO systems. The authors claims that the proposed solution is having relatively lower complexity.

In [19] authors presented an extensive review on millimeter wave-based Massive MIMO system and emphasized on the associated technologies, emerging trends, applications, associated challenges, and future directions. This survey article covered the technology from root basics until the recent advancements. In [20], authors presented the expected performance parameter from 5G networks then the interconnection between densification of networks, millimeter wave communication and Massive MIMO systems is beautifully presented.

In [21], the authors presented the journey of communication systems from SISO to Massive MIMO system with their relative advantages and associated challenges.

Furthermore, the authors presented the propagation characteristics of millimeter wave spectrum and Massive MIMO architectures. In addition, the variety of precoding techniques for different Massive MIMO configurations are discussed and compared. In [22] authors covered almost every aspect of future wireless communication systems including health issues and at the last authors concluded that the millimeter wave Massive MIMO system have a very vast potential to fulfill the requirements of future wireless communication systems.

In [23] authors presented an extensive review of Massive MIMO system under the influence of channel and hardware impairments. The authors highlighted the fact that in MIMO systems, enhancement in spectral efficiency is achieved by reusing the frequency as well as time slots for multiple users in a single cell. This has associated challenges that can degrade the system performance. In [24], authors mainly focused on channel and hardware impairments like channel estimation, fast moving user equipment, and channel aging. Which can significantly degrade the system performance. Further, the authors compared system performance considering various communication scenarios, highlighted the significant work and provided the future directions.

In [25] authors explored the requirements, trend, associated technologies, and application of modern Internet of Things (IOT) systems. The authors pointed that the upcoming 5th generation communication technology is a key enabler for modern IOT systems. The authors present the emerging technologies for IOT, mainly 5G and also covered the challenges and research opportunities in IOT applications.

3 Design Methodology

From the rigorous literature review presented in previous section, it is evident that implementation of Massive MIMO systems with hybrid channel estimation techniques in millimeter wave spectrum will lead to significant performance enhancements. Researchers around the globe are continuously working to address the multiple challenges before actual implementation of the technology. Following are the main challenges encountered and needed to be addressed before implementation of the same.

- Improving Spectral Efficiency of Hybrid Architecture.
- Reducing Hardware Complexity of Hybrid Architecture.
- Reducing Computational Complexity of Hybrid Channel Estimation.
- Reducing the High-Power Consumption.
- Maintaining Performance under Different Channel Conditions.

The above-mentioned challenges are the main interferences in implementation of Massive MIMO systems in millimeter wave spectrum. Extensive research must be carried in order to improve the overall performance of Massive MIMO systems w.r.t. key metrics like spectral efficiency, hardware efficiency, and energy efficiency.

Keeping in mind the above stated points, this work is targeted at achieving the near optimal spectral efficiency with reduced hardware/computational complexity of hybrid channel estimation Massive MIMO systems. It allows for multiplexing gains utilizing DLCNN estimation by connecting many user equipment's to the corresponding time–frequency service. Furthermore, it has more base station antennas than user equipment per cell, as seen in Fig. 2, to effectively eliminate interference. The base stations should intensify the number of antennas proportionally if the predicted number of user equipment's in a cell increase. Time Division Duplex (TDD) mode reduces the overhead CSI recovery induced by many antennas and does not rely on channel models. The large-scale MIMO incorporates these architecture principles to make high SE in the coverage level of future wireless systems an effective way. Each base station is fitted with a wide variety of antennas, M, and serves a cell, K. Each terminal usually has a single antenna. Various bases represent different cells. With the possible limitation of power management and pilot allocations, Massive MIMO does not depend on base station cooperation. In both uplink and downlink transmissions, all terminals have used the entire time–frequency capabilities simultaneously. The multiplexing and de-multiplexing signal processing of the base station is provided with multiplicative antennas and CSI due to TDD methods, calculating the terminals' pilots and reciprocating the uplink and downlink transmissions. The CSI needs transceiver hardware reciprocity calibration.

The Frequency Division Duplex (FDD) Downlink Training with the appropriate CSI feedback provides broad M and K, an inadmissible high uplink overhead, considering the higher multiplexing advantage. The downlink channel from uplink training is estimated using TDD systems with channel reciprocity. It is noteworthy both in terms of achieving performance and simplifying downlink scheduling at BS. The transmitter knowledge of previous channel outcomes acquired by training or channel statistics can be determined based on a short sequence. However, additional feedback is required to obtain channel statistics on massive non-stationary MIMO channels. Thus, TDD/FDD large MIMO systems are needed with a different CSI estimate and feedback methodology that provides reliable CSI with low overheads and low complexity. To address this challenge, DLCNN channel estimation is a palatable technique for estimating a sparse channel in uncertain channel statistics.

3.1 DLCNN Channel Estimation

Due to its amplitude, phase, and wavelength, a wireless communication channel has significant random ratings. Measurement of the channels and recognition of reception as a receiver's decisive components is essential for wireless communications. Communication systems usually complete channel estimation at the receiver, and the precision of the estimation has a significant effect on overall device performance. Channel prediction algorithms can be classified into two types based on the transmitting signal: frequency domain and time domain. They can be known as reference signal estimates, blind estimates, or half-blind estimates depending upon whether

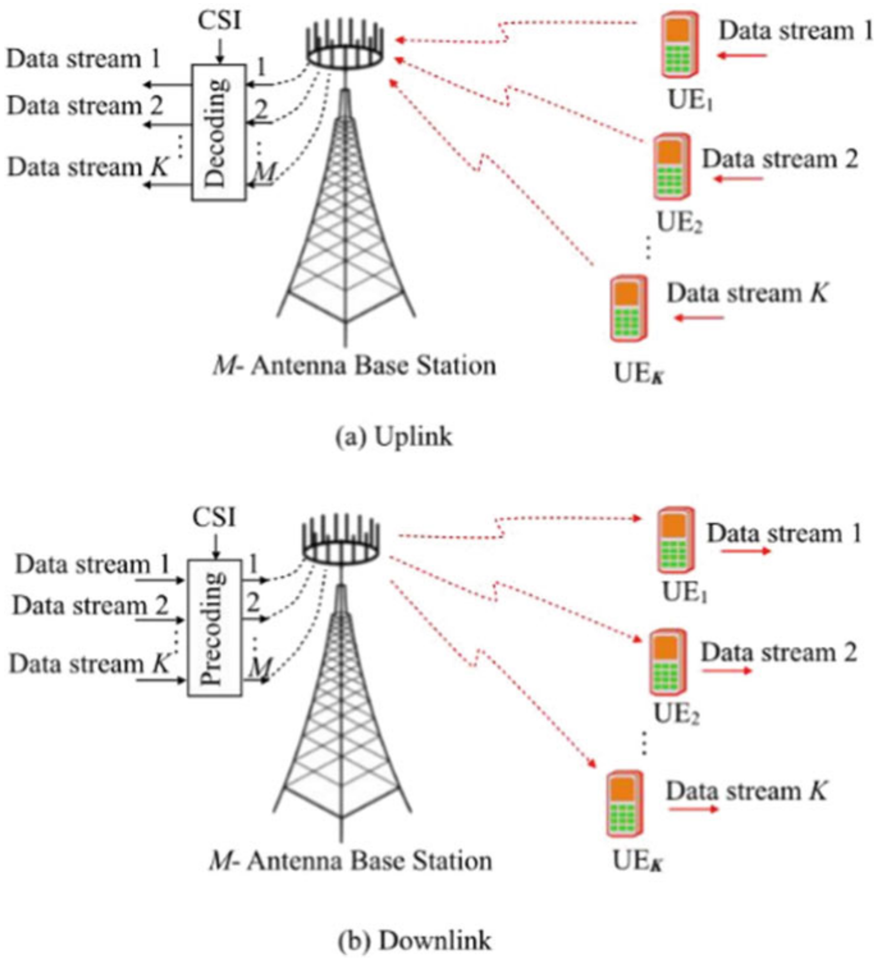


Fig. 2 Massive MIMO system model

prior information is required. Usually, the pilot sequence provides a more convincing approximation result based on the reference signal. In order to achieve high spectral efficiencies for blind and half-blind estimates, a brief sequence is needed. Nevertheless, the associated drawbacks are poor precision of estimates, significant uncertainty in the process, and estimated non-convergence issues. To increase the bandwidth of the channel, the transmitter and receiver need the CSI. Figure 3 shows the DLCNN-based channel estimation, where CSI of the transmitter is recovered and transmitted through the receiver. In the practical application of Massive MIMO technology, reception identification of the Massive MIMO communication device is critical. There are two types of signal detection algorithms: linear and nonlinear. The obtained signal x will reconstruct the initially transmitted signal entirely through linear operations

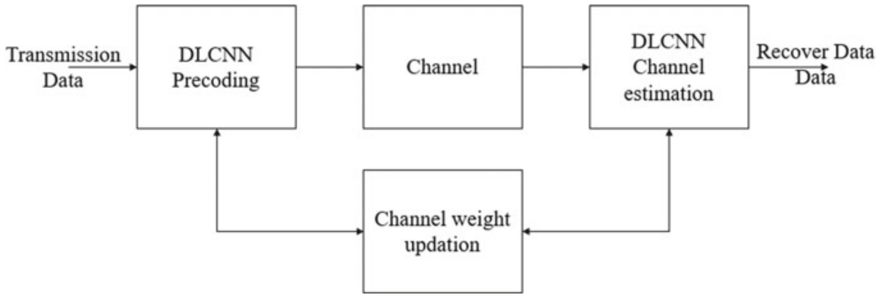


Fig. 3 DLCNN channel estimation

in linear detection

$$y = h_{\text{llk}}^{\text{prop}} x + z \tag{1}$$

Here, $h_{\text{llk}}^{\text{prop}}$ is the recipient's computing matrix generated from DLCNN model. The Massive MIMO technique is extensively practiced in mobile wireless networks. It will affect the next generation's wireless networking because it transmits rapidly and has a high bandwidth capacity that is likely to decrease and slow multipath. The Massive MIMO signal structure performs it very difficult for channel calculations to match unsafety and efficacy. The pilots, who seem to be both the transmitter and the destination's reference signal, are practiced in several channels estimating schemes suggested in the research. The design of pilots and an effective estimation algorithm are the two demanding problems for this solution. Many pilots use accurate calculations to determine the majority of existing pilot-assisted channel forecasts; hence the spectral efficiency is lessened.

Figure 4 illustrates the layer architecture of DLCNN model. The pilot signal and guard band are paired in this arrangement with a modulated signal in the frequency domain of $X(k)$, $k \in [1, N]$. The signal is then translated into the time domain by N-point IDFT denoting $X(n)$, $n \in [1, N]$ and is used to eliminate inter-symbol and inter sub-carrier interferences by a cyclic extension of the length. The resultant time series data transformed into an analog transmission signal by employing a DAC conversion at1/TS Hz clock speed. The channel response generated from DLCNN model contains a propagation path modeled with P taps in an intricate time domain vector.

$$h_{\text{llk}}^{\text{prop}} = \sum_{p=1}^{P} \alpha_p \delta(n - T_P T_S) \tag{2}$$

Here, T_P is the multipath delay $(0 \leq T_P T_S \leq T_G)$ and is the specific multipath portion. Since, $T_G < T_S$, so, at first, the non-zero channel response concentrates, which illustrates to $h = [h_1, h_2, \ldots, h_n]$; i.e., the first N coefficients of h are the

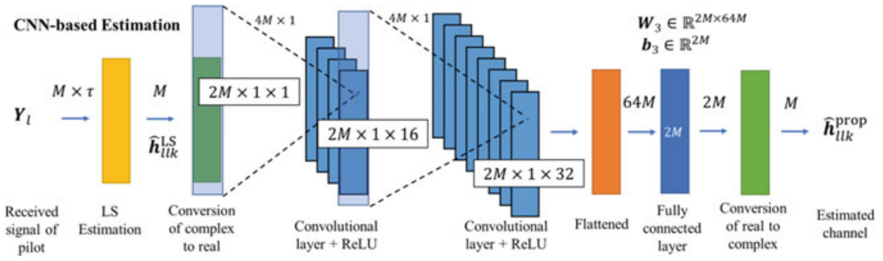


Fig. 4 Layered architecture of DLCNNs

only once that can have non-zero elements. Considering that all interferences are excluded, the relationship between channel response and noise is described by

$z(n)$ is:

$$z = xh^{\text{prop}}_{\text{llk}} \tag{3}$$

where ⊗ indicates convolution and $\xi(n) \in [1, N]$ signifies the Gaussian noise.

4 Results and Discussion

The simulations are carried out using the MatllabR2022a platform and simulation platform like energy efficiency, spectrum efficiency is compared with state-of-art approaches. The UE density is the input parameter, which shows the available number of UE in the square kilometer area. In addition, BS density is another input parameter, which shows the available number of BS in the square kilometer area.

Figure 5 compares the spectrum efficiency performance of various channel estimation methods in Massive MIMO environment for multiple number of UE density. Here, the existing LS [19] method resulted in 3.2 Mbits/UE, existing MMSE [14] method resulted in 8.2 Mbits/UE, and proposed DLCNN channel estimation resulted in 10.2 Mbits/UE. Therefore, the proposed DLCNN channel estimation resulted in higher spectrum efficiency performance comparison as compared to LS [19] and MMSE [14] methods. Figure 6 compares the energy efficiency performance of various channel estimation methods in Massive MIMO environment for multiple number of BS densities. Here, the existing LS [19] method resulted in 7.8 Mbits/Joule, existing MMSE [14] method resulted in 10.4 Mbits/joule, and proposed DLCNN channel estimation resulted in 16 Mbits/Joule. Therefore, the proposed DLCNN channel estimation resulted in higher energy efficiency performance comparison as compared to LS [19] and MMSE [14] methods. Figure 7 shows the need of number of BSs (i.e., BS density) for multiple number of UEs (i.e., UE density). Here, the proposed DLCNN channel requires the lesser number of BSs as compared to MMSE [14], and LS [19] methods.

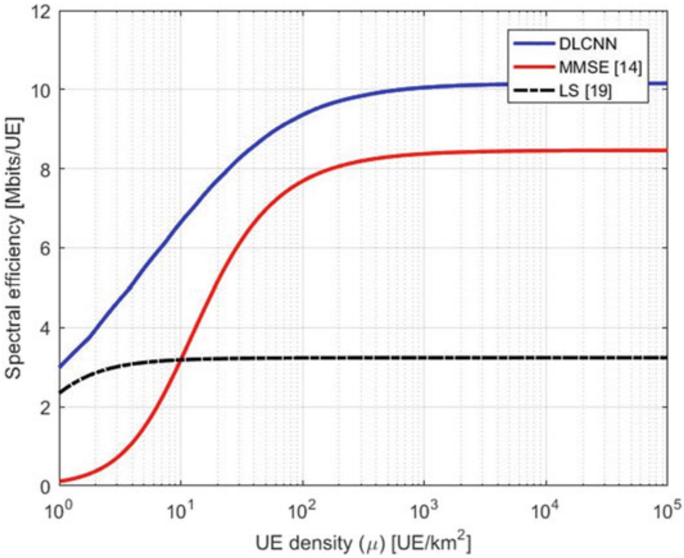


Fig. 5 Spectrum efficiency comparison

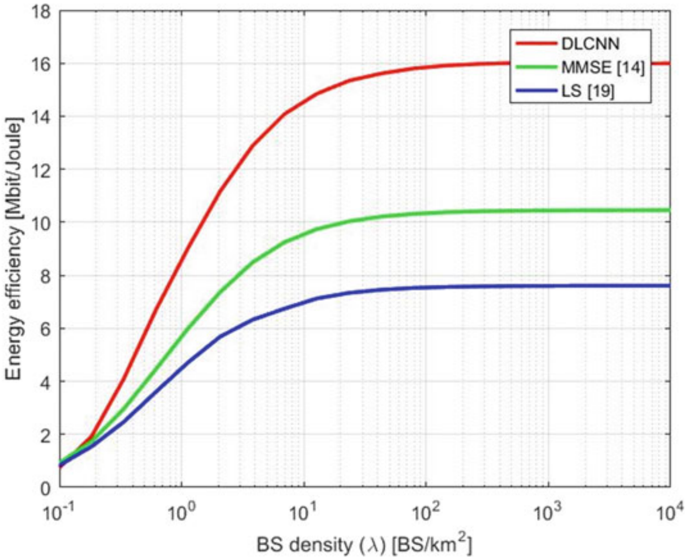


Fig. 6 Energy efficiency comparison

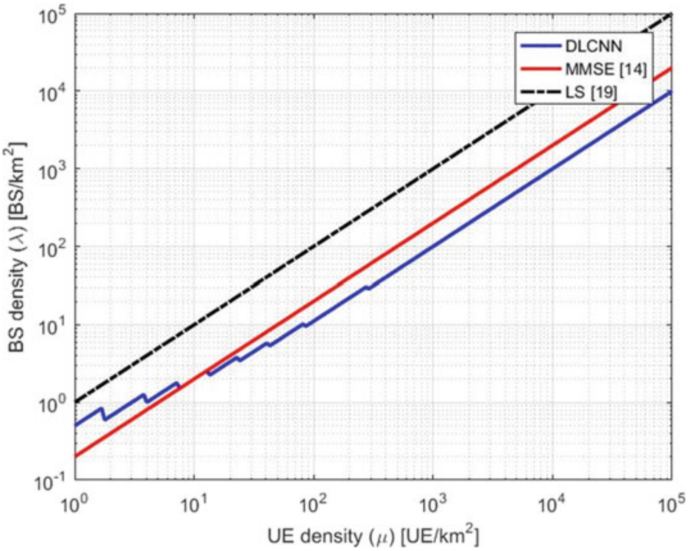


Fig. 7 BS density versus UE density

5 Conclusion

This article is focused on analysis of DLCNN channel estimation for Massive MIMO environment. Usually, the uplink and downlink communications of Massive MIMO were suffering with higher uncertainties and noises. Therefore, the channel estimation needs auto updated coefficients based on noise properties, which was accomplished by adopting the artificial intelligence approach. Here, DLCNN model was trained with the various noise conditions under fading channel properties, which can effectively estimate received data with error free properties. The simulation results showed that the proposed DLCNN channel estimation resulted in superior spectrum efficiency, energy efficiency performance as compared to state-of-art approaches. Further, this work can extend with extreme learning machine-based hybrid precoding and beam forming for reduced error rates.

References

1. Pirzadeh H, Swindlehurst AL (2018) Spectral efficiency of mixed-ADC massive MIMO. IEEE Trans Sig Process 66(13):3599–3613
2. Özdogan Ö, Bjöornson E, Zhang J (2018) Cell-free massive MIMO with Rician fading: estimation schemes and spectral efficiency. In: 2018 52nd Asilomar conference on signals, systems, and computers. IEEE
3. Shlezinger N, Eldar YC (2018) On the spectral efficiency of noncooperative uplink massive MIMO systems. IEEE Trans Commun 67(3):1956–1971

4. Zhang M et al (2018) Spectral efficiency and power allocation for mixed-ADC massive MIMO system. China Commun 15(3):112–127
5. Björnson E, Sanguinetti L, Hoydis J (2018) Hardware distortion correlation has negligible impact on UL massive MIMO spectral efficiency. IEEE Trans Commun 67(2):1085–1098
6. Liu P et al (2019) Spectral efficiency analysis of cell-free massive MIMO systems with zero-forcing detector. IEEE Trans Wirel Commun 19(2):795–807
7. Salh A et al (2019) Trade-off energy and spectral efficiency in a downlink massive MIMO system. Wirel Pers Commun 106(2):897–910
8. Hei Y et al (2019) Energy and spectral efficiency tradeoff in massive MIMO systems with multi-objective adaptive genetic algorithm. Soft Comput 23(16):7163–7179
9. Du J et al (2019) Weighted spectral efficiency optimization for hybrid beamforming in multiuser massive MIMO-OFDM systems. IEEE Trans Veh Technol 68(10):9698–9712
10. Liu G et al (2019) Joint pilot allocation and power control to enhance max-min spectral efficiency in TDD massive MIMO systems. IEEE Access 7:149191–149201
11. Mai TC, Ngo HQ, Duong TQ (2020) Downlink spectral efficiency of cell-free massive MIMO systems with multi-antenna users. IEEE Trans Commun 68(8):4803–4815
12. You L et al (2020) Spectral efficiency and energy efficiency tradeoff in massive MIMO downlink transmission with statistical CSIT. IEEE Trans Sig Process 68:2645–2659
13. Pirzadeh H et al (2020) Spectral efficiency of one-bit sigma-delta massive MIMO. IEEE J Sel Areas Commun 38(9):2215–2226
14. Nguyen TH et al (2021) Pilot assignment for joint uplink-downlink spectral efficiency enhancement in massive MIMO systems with spatial correlation. IEEE Trans Veh Technol 70(8):8292–8297
15. Arshad J et al (2020) Spectral efficiency augmentation in uplink massive MIMO systems by increasing transmit power and uniform linear array gain. Sensors 20(17):4982
16. Jin S-N, Yue D-W, Nguyen HH (2020) Spectral efficiency of a frequency-selective cell-free massive MIMO system with phase noise. IEEE Wirel Commun Lett 10(3):483–487
17. Jiang B et al (2020) Energy efficiency and spectral efficiency tradeoff in massive MIMO multicast transmission with statistical CSI. Entropy 22(9):1045
18. Dicandia FA, Genovesi S (2021) Exploitation of triangular lattice arrays for improved spectral efficiency in massive MIMO 5G systems. IEEE Access 9:17530–17543
19. Chen L, Zhang L (2021) Spectral efficiency analysis for massive MIMO system under QoS constraint: an effective capacity perspective. Mobile Netw Appl 26(2):691–699
20. Salh A et al (2021) Trade-off energy and spectral efficiency in 5G massive MIMO system. arXiv preprint arXiv:2105.10722
21. Dicandia FA, Genovesi S (2021) Spectral efficiency improvement of 5G massive MIMO systems for high-altitude platform stations by using triangular lattice arrays. Sensors 21(9):3202
22. Zhang X et al (2021) Spectral efficiency improvement and power control optimization of massive MIMO networks. IEEE Access 9:11523–11532
23. Amadid J et al (2022) On channel estimation and spectral efficiency for cell-free massive MIMO with multi-antenna access points considering spatially correlated channels. Trans Emerg Telecommun Technol:e4438
24. Mirhosseini FS, Tadaion A, Razavizadeh SM (2021) Spectral efficiency of dense multicell massive MIMO networks in spatially correlated channels. IEEE Trans Veh Technol 70(2):1307–1316
25. Zhang Y et al (2021) Spectral efficiency of superimposed pilots in cell-free massive MIMO systems with hardware impairments. China Commun 18(6):146–161

A Review on Sign Language Recognition Techniques

S. Rakesh, M. Venu Gopalachari, D. Jayaram, Ishika Gupta, Kritika Agarwal, and Ganji Nishanth

Abstract Not many people understand sign language, which poses a problem to people who can only communicate with sign language. They are unable to perform basic day-to-day activities due to this problem. Thus, many solutions were developed to help such people communicate effectively and live relatively easier life. Many papers from reputed journals were studied to gain an understanding of the working of these solutions. Based on our research, there are two kinds of solutions proposed: a glove-based approach where mute people can wear the gloves and sign what they want to communicate, and the device placed on the glove can convert the gestures signed to speech or text, which a non-mute person can easily understand and comprehend. The second solution is a computer vision based. By using a camera, the gestures signed by a person can be captured, processed, and identified. The output can be given in the form of text or speech or both. The rest of this review paper goes in depth about the various solutions proposed. Their methodology, advantages, and disadvantages are analyzed, compared, and discussed.

Keywords Deep learning · Sign language · Sign to speech

1 Introduction

Most of us in today's society do not realize how powerful words can be and how huge of an impact they can have on our lives. Effective communication helps us in expressing our thoughts, ideas, and feelings. It helps us build lasting relationships in both personal and professional life. Unfortunately, not many people have the privilege to easily communicate with each other using speech. There are people who are unable to speak. They cannot use their voice to articulate words. Thus, sign language was created to enable such people to communicate. Sign language is a medium of communication using visual hand gestures called signs. There are

S. Rakesh (✉) · M. Venu Gopalachari · D. Jayaram · I. Gupta · K. Agarwal · G. Nishanth
IT Department, Chaitanya Bharathi Institute of Technology, Hyderabad, India
e-mail: srakesh_it@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_32

more than 70 million people in the world, who use sign language. But there is not a single universal sign language that people use to communicate. There are more than 300 sign languages that are used across the world. They can differ from country to country. Even within the same country, sign languages can have subtle differences.

2 Literature Survey

In reference [1], the authors of the paper have proposed an algorithm that uses YCbCr color space, COG, and template matching for segmenting the hand, detecting the fingers, and then identifying the hand gesture from real imagery to help the deaf and mute people. To discover isolated indicators, the authors of [2] use two appearance-based algorithms, I3D and TimeSformer, and one posed base approach, SPOTER, on two publically available datasets, AUTSL and WLASL300. They increased the accuracy of the WLASL300 dataset by applying the CMA-ES optimization technique. They have also proposed neural ensembler, an ensemble approach based on a transformer model. The authors of [3] used techniques such as fast accelerated segment test (FAST), scale-invariant feature transformation (SIFT), and convolution neural networks (CNN) [4–6] for effective feature extraction, which is required for automatic gesture detection. FAST and SIFT are used to recognize and calculate features from pictures, respectively. CNN was utilized for classification, with FAST-SIFT characteristics hybridized. Vedak et al. [7] described a method for capturing a still hand image frame with a webcam. These frames were enhanced by post-processing. The sign language is then translated into English text using feature extraction and classification techniques [8, 9]. The text to speech API is used to transform this translation to speech. Pathak et al. [10] used a user-defined dataset with around 2000 photos, approximately 400 for each of its classes, to develop and implement a sign language recognition model based on a CNN. To train the dataset, they used a pre-trained SSD MobileNet V2 architecture.

The authors of [11] used transfer learning and deep CNN [12–16] fine-tuning to recognize 32 Arabic sign language hand movements. The networks were fed regular 2D images of various Arabic sign language data. Thap et al. [17] used blob analysis to extract the hand gesture from the image considering that the largest blob is the hand. Then the blob was resized and sampled, and length encoding was done. To identify the hand signs, the authors of [18] took a different approach. It employs algorithms (Gauss Laplace algorithm) and methods for performing sign language translation. The vision of an efficient system for translating sign language to text is well within reach, but the challenges lie in optimization. A colored tape is applied to the fingers to serve as an input. We must show these fingers to a computer that has MATLAB installed. In [19] this paper, sign language recognition is done using convolutional neural network. This system is able to interpret static hand gestures. This system converts uploaded video to text and then converts that text to a gif. This model uses dense layer and softmax regression in the classification stage in CNN implementation.

3 Methodology

According to the methodology proposed by Gupta et al. [1], the user will have to upload their video of performing a sign. The video is then divided into frames, and each frame is used as an image. To segment the hand from the background, the RCB image is converted to YCbCr color space. The centroid of gravity (COG) is then calculated. The farthest distance from the centroid to the tip of the longest active finger in the particular gesture is taken as the radius to draw a circle. The circle consists of the gesture. Thus, COG was used to segment the hand. The unknown gesture is then compared with the preset template models of individual gestures to identify the gesture (Fig. 1).

Hrúz et al. [2] implemented I3D (a family of convolutional neural network models for video classification) with three inputs: Crop and Resize, Masked, and OptFlow. As input, 16 frames from each video were taken. The Kinetics400 dataset was used to pre-train all I3D models. The authors followed the original Facebooksearch GitHub implementation. The authors implemented sign pose-based transformer (SPOTER) on top of body-pose sequences using PyTorch. They followed Boháek M., Hrz M. sign pose-based transformer for word-level sign language recognition training. The SGD optimizer with a learning rate of 0.001 was used. Momentum and decay were both set to zero. The initial weights of the model were drawn from a uniform distribution of [0, 1]. On WLASL300, they trained SPOTER for 100 epochs. They trained SPOTER with OpenPose and MMPose on the AUTSL dataset for 12 and 35 epochs,

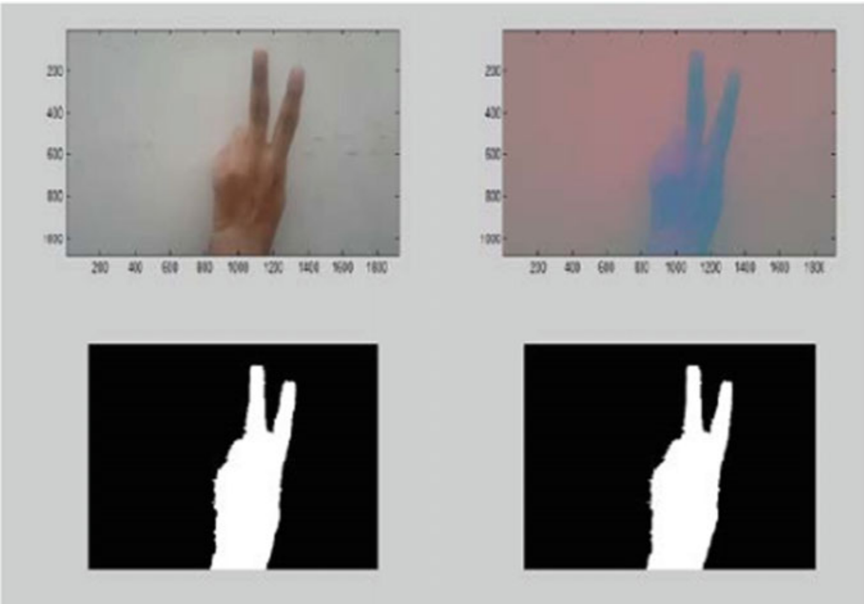


Fig. 1 Image conversion

Fig. 2 Different data modalities are given as input

respectively, until each model variant converged on the training split. The authors used a neural ensemble to combine the outputs of individual models and make a final decision. PyTorch was used to implement it (Fig. 2).

FAST is a corner detection method that is used to extract feature points, which are then used in many computer vision tasks such as tracking and mapping. SIFT is a computer vision algorithm that detects, describes, and matches local features in images while remaining independent of image scale and rotation. In [3], authors proposed three major phases for sign recognition: data preprocessing, feature extraction, and CNN training and testing. The stored static single-handed images are resized in the first phase, and then data augmentation is performed on those resized images. FAST techniques are then used to localize key points. In the third phase, the value of these localized key points was calculated using SIFT. These values were then fed into CNN for training and classification. Keras, a Python-based library, was used to implement and test the system.

The different steps described in [7] to detect signs are first, the hand gestures are captured by video, and they are then divided into frames. Image preprocessing is then done, and then feature extraction is performed using Histogram of Oriented Gradients (HOG). Support vector machine (SVM) is used for classification and recognition. The text is given as the output. The Google API is used for converting text to speech to give audio as the output as well. The authors have created their own dataset of 26 English alphabets of ISL. Four thousand and eight hundred images were used for training, and 1200 images were used for testing.

In paper [10], the three steps were recognized to design the model—obtaining footage of the user signing and taking them as input, classifying each frame in the video as a sign, and reconstructing and displaying the most likely sign from classification scores to provide that as output. They made use of a convolutional neural network, a deep learning system that can take an input image and assign importance to various elements and objects in the image, as well as distinguish between them. When compared to other classification systems, a ConvNet requires far less preprocessing. ConvNets can pick up these filters and properties with the right training, but basic approaches require filter engineering by hand. A dataset of 2000 images with 400 images divided among five classes was used to train this model.

The suggested method in [17] makes use of a skin color model to recognize the hand in the image, and additional preprocessing is carried out to eliminate unnecessary noise and regions. Since the hand is the biggest blob in the image, blob analysis is used to extract the hand gesture from the picture. Then, in order to remove the size variance limitation, the blob is shrunk to a standard size. The junction point between the grid line and the boundary is then extracted. The hand gesture's border is represented using the Freeman chain coding. In order to reduce the duration of encoding for chain code, run length is completed. Its shape number is discovered by locating the first variation in the chain code. Each motion may be uniquely identified using a number (Fig. 3).

In [18], a colorful tape is wrapped around the fingers and used as an input. We must present these fingers to a computer that has MATLAB installed. The output of the computer is translated to digital and processed by the microcontroller before responding as voice via the speaker. To generate the output in this project, we used a microcontroller, a voice IC, and a speaker. Power supply and voice IC are employed as hardware components (aPR33a3). MATLAB is the software utilized.

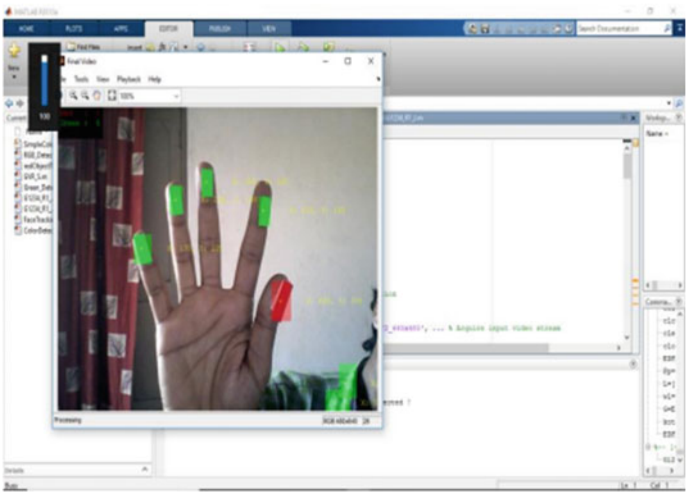


Fig. 3 Gauss Laplace edge detection

4 Result Analysis

The major advantage of [1] was that their algorithm was able to detect the hand gesture from any background (robust to background) and was able to identify the hands of different skin tones, thereby giving a better result. But at the same time, it was computationally expensive and could not detect sign language in real time, since the user would have to provide a video of them.

The individual models implemented in [2] gave suboptimal results. Appearance-based approaches were able to distinguish two similar signs, but it required a larger training dataset. Pose-based approach required a smaller training dataset but was unable to distinguish between similar signs. By using ensemble techniques, better accuracies were achieved: 96.37% in AUTSL dataset and 73.87% in the WLASL300 dataset. The neural ensemble method proposed by the authors has high potential for future research. But isolated sign language recognition still poses a problem in an uncontrolled environment.

The findings of the methodologies suggested in [3] were tested on 24 alphabet sets and ten-digit sets, 34 Indian sign language motions, and two publicly accessible datasets, the Jochen Trisech Dataset (JTD) and the NUS-II dataset. The proposed FiST CNN model outperformed previous models such as CNN and SIFT CNN. In terms of accuracy and calculation time, FiST CNN outperformed. The FiST CNN attained accuracy of 97.89%, 95.68%, 94.90%, and 95.87% for ISL-alphabets, MNIST, JTD, and NUS-II, respectively. The model cannot detect word-level signals, which is a drawback (Fig. 4).

The overall accuracy of the model proposed by Vedak et al. [7] was 88%. But this application is not portable. It could be made into a mobile application for convenience. Also, it cannot identify word-level signs. The same drawback is mentioned in [3].

The proposed system in [10] gave accurate results under controlled light and intensity. As a result, by extending the dataset, the model may be simply extended on a

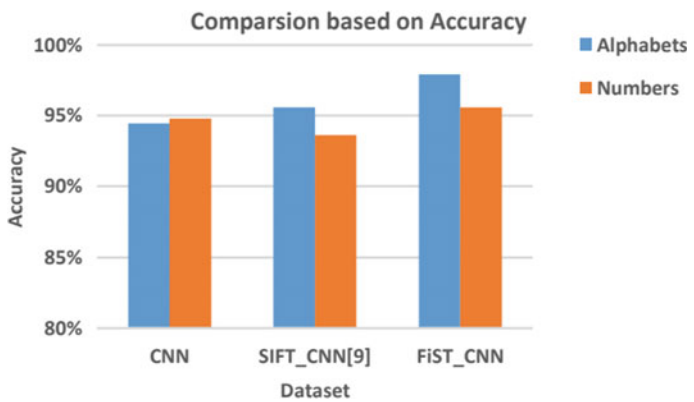


Fig. 4 Accuracy comparison

wide scale. However, this approach has significant drawbacks, such as environmental conditions such as low light intensity and unmanaged backdrop that reduced detection accuracy.

By converting the gesture to a shape number, the object descriptor used in [17] to convey boundary information made categorization simple and straightforward. After analyzing a few NSL movements, this method for gesture detection is determined to be straightforward and satisfactory. It still has to be tested on additional gestures, though.

In [18], despite using a low-quality camera, this system was able to process them and produce a video output. All input images were captured by webcam or an android device. To minimize the time and memory requirements associated with pattern recognition, the captured image is converted into binary form after reading into MATLAB. The equivalent image is then shown using PCA, which calculates the Eigen vectors with the minimum equivalent distance for the input gesture. After looking up the input image in the database, it provides the corresponding hand gesture and alphabet. The computation may differ highly based on the type of background in images.

5 Conclusion

From the above-mentioned information, it can be concluded that there are several solutions proposed to help the mute people communicate. The two main types of solutions offered are glove based and computer vision based. While it is a breakthrough that signs can now be converted to speech or text by using various machine learning and deep learning techniques, there are still some challenges to overcome. Some of the most common disadvantages found are that most of the solutions work well with just numbers and letters. The person who is signing would have to spell out all the words to be able to communicate a sentence. This is not a way of realistic communication. Moreover, a sophisticated design has not been developed yet for the glove-based approach [20–23]. It would be very inconvenient for the person to wear a glove with a microcontroller placed on top of it, as it would neither be comfortable nor portable. Thus, we propose a solution where we build a mobile app that can convert sign language to speech (in the language preference of the user) by capturing the signs via camera and then identifying the signs real time by using deep learning techniques [24–27]. This will ensure a more realistic communication between the mute and non-mute people and could also be easily used by people, since it would be an app.

References

1. Gupta H, Ramjiwal A, Jose JT (2018) Vision based approach to sign language recognition. Int J Adv Appl Sci (IJAAS) 7(2)
2. Hrúz M, Gruber I, Kanis J, Boháček M, Hlaváč M, Krňoul Z (2022) One model is not enough: ensembles for isolated sign language recognition. PubMed Central
3. Tyagi A, Bansal S (2022) Hybrid FiST-CNN approach for feature extraction for vision-based Indian sign language recognition. Int Arab J Inf Technol 19(3)
4. Shahin OR, Taloba AI, El Rwelli R (2021) Gesture based Arabic sign language recognition for impaired people based on convolution neural network. (IJACSA) Int J Adv Comp Sci Appl 12(12)
5. Aldhahri E, Aljuhani R, Alfaidi A, Alshehri B, Alwadei H, Aljojo N, Alshutayri A, Almazroi A (2022) Arabic sign language recognition using convolutional neural network and MobileNet. Arab J Sci Eng
6. Khosravani Pour L, Farrokhi A (2022) Language recognition by convolutional neural networks. Scientia Iranica
7. Vedak O, Zavre P, Todkar A, Patil M (2019) Sign language interpreter using image processing and machine learning. Int Res J Eng Technol (IRJET) 6(4)
8. Tamiru NK, Tekeba M, Salau AO (2022) Recognition of Amharic sign language with Amharic alphabet signs using ANN and SVM. Springer
9. Gowtham V, Karthick S, Karthikeyan T, Elayaraja P Dr (2022) Sign language recognition using machine intelligence for hearing impairment person. IJERT 10(9)
10. Pathak A, Kumar A, Priyam A, Gupta P, Chugh G (2022) Real Time sign language detection. Int J Mod Trends Sci Technol
11. Saleh Y, Issa GF (2020) Arabic sign language recognition through deep neural networks fine-tuning. Int J Online Biomed Eng (iJOE) 16(5)
12. Ravi Kumar A Dr, Bhavana T, Sri PM (2022) A deep neural framework for continuous sign language recognition by iterative training. J Algebr Stat 13(3):4574–4584
13. John J, Sherif BV (2022) Hand landmark-based sign language recognition using deep learning. Springer
14. Rastgoo R, Kiani K, Escalera S (2022) Real-time isolated hand sign language recognition using deep networks and SVD. Springer
15. Sharma S, Singh S (2021) Recognition of Indian sign language (ISL) using deep learning model. Springer
16. Tanseem NA, Abu-Naser SS (2022) Classification of sign-language using MobileNet—deep learning. Int J Acad Inf Syst Res (IJAISR) 6(7):29–40
17. Thap V, Sunuwar J, Pradhan R (2019) Finger spelling recognition for Nepali sign language. JCSE Int J Comp Sci Eng 3(6)
18. Sruthi R, Venkateswara Rao B, Nagapravallika P, Harikrishna G, Narendra Babu K (2018) Vision based sign language by using MATLAB. Int Res J Eng Technol (IRJET) 5(3)
19. Rakesh S, Bharadhwaj A, Harsha ES. Sign language recognition using convolutional neural network
20. Hasan W, Gabeal NN (2019) Implementation smart gloves for deaf and dumb disabled. Int Sci Technol J 19
21. Sanjay S, Dinesh Kumar S, Suresh M Dr, Vasantha Kumar S, Saravana Kumar K, Mohamed Rifayee Hussain Z (2021) Development of smart glove for deaf-mute people. Nat Volatiles & Essent Oils J
22. Bhore PS, Londhe ND, Narsale S, Patel PS, Thanambir D (2020) Smart gloves for deaf and dumb students. Int J Progr Res Sci Eng 1(8)
23. Joksimoski B, Zdravevski E, Lameski P, Pires IM, Melero FJ, Puebla T (2022) Technological solutions for sign language recognition: a scoping review of research trends, challenges, and opportunities. IEEE
24. Zhou Z, Tam VWL, Lam EY (2022) A cross-attention BERT-based framework for continuous sign language recognition. IEEE Sig Process 29

25. Montefalcon MD, Padilla JR, Rodriguez R (2022) Filipino sign language recognition using long short-term memory and residual network architecture. Springer
26. Subramanian B, Olimov B, Naik SM, Kim S, Park K-H, Kim J (2022) An integrated media pipe-optimized GRU model for Indian sign language recognition. Springer
27. Inamdar N, Inamdar Z, Sheikh S, Sushir SA (2022) A survey paper on sign language recognition. Int J Res Appl Sci Eng Technol (IJRASET) 10(IV)
28. Mohammedali AH, Abbas HH, Shahadi HI (2022) Real-time sign language recognition system. Int J Health Sci 6(S4):10384–10407
29. Ghotkar A, Gokhale A, Malpure D (2022) Real-time left- and right-hand detection for sign language recognition. Springer
30. Bilge YC, Cinbis RG, Ikizler-Cinbis N (2022) Towards zero-shot sign language recognition. IEEE

Review on 3D Model Generation Through Natural Language and Image Processing

Sri Chakra Raj Pyaraka, Sravan Sai Lanka, Pranav Konanki, M. Venu Gopalachari, S. Rakesh, and D. Jayaram

Abstract Extended reality also known as XR is the future of technology, and augmented reality sits right in the front. Every aspect of the IT industry is slowly embedding AR into its interfaces enabling more interactivity for users. Augmented reality (AR) is the integration of digital information with the user's environment in real time. Augmented reality is hugely dependent on 3D models apart from audio visuals. Developing 3D models for every aspect possible consumes a lot of time and hard work. Automating this repetitive work and manual effort will save a lot of time and energy. The models generated can be used in both augmented and virtual realities. Work in text-to-image generation has been significant, but image-to-model generation is a tedious task involving referencing several images, detecting the dimensions of the desired object, and processing it for model generation. Our efforts were into detecting the edges of objects in input images and automating the task of building a model using data obtained from the processed image.

Keywords 3D models · Augmented reality · XR-extended reality · Image-to-3D

1 Introduction

Automaton involves the automation of image-to-model generation with the help of image processing techniques, and Python interfaces to communicate with 3D modeling software. Image processing involves five step process to ensure noise is reduced, and outlines of desired objects are determined through a modified Sobel operator. Edge-detection techniques can be applied to detect linear segments in an image, categorized into gradient and Gaussian domains. The outline data is sent to a 3D modeling software through Python API. After the model is generated, it goes through a series of evaluation steps to ensure that model is almost accurate as depicted in the picture, else parameters of image processing steps are changed to meet the

S. C. R. Pyaraka · S. S. Lanka · P. Konanki · M. Venu Gopalachari (✉) · S. Rakesh · D. Jayaram
IT Department, Chaitanya Bharathi Institute of Technology, Hyderabad, India
e-mail: mvenugopalachari_it@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_33

requirements. The model output will be made available for viewing online as well as a downloadable 3D file.

2 Literature Review

In this section, we will briefly review the related works done in the field of 3D model generation and related work regarding image processing via edge detection. Human perceptual evaluation helps us to evaluate whether the view-generated model is like the respective source image. We can automate this human evaluation procedure with the help of machine learning. The text-to-model generation helped us understand the solution to generate 3D models from different views. This paper made it possible to build 3D models from the text.

The authors have followed the following principles while developing the algorithm:

(a) Modular framework
(b) Lower computations are preferred over a small loss of accuracy
(c) Better edges are preferred over lower noise.

3 Methodologies and Results

3.1 Text-to-3D Generation

Significant methodologies were found in text to 3D generation which involved natural language processing and image processing techniques to develop better 3D models. Diffusion models trained on billions of image-text pairs have been the driving force behind recent advances in text-to-image synthesis [1]. Large-scale datasets of labeled 3D data and effective designs for 3D data de-noising are necessary for adapting this method for 3D synthesis. In the recent study, execution of text-to-3D synthesis utilizing a pre-trained 2D text-to-image diffusion model helped in building the requirements. It describes a loss based on probability density distillation that makes it possible to optimize a parametric picture generator using a 2D diffusion model as a prior. Optimizing a randomly initialized 3D model (NeRF—neural radiance field) via gradient descent using this loss in a method akin to DeepDream so that the 2D renderings from various angles have less loss.

NeRF-like models have been successfully incorporated as building blocks within bigger generative systems by numerous 3D generative methodologies [1]. One such method is Dream Fields, which trains NeRFs using an optimization-based method and employs frozen image-text joint embedding models from CLIP. This work demonstrated the use of pre-trained 2D image-text models for 3D synthesis, albeit the 3D objects created by this method typically lack precision and realism. Other methods based on voxel grids and meshed have been guided by CLIP [2].

It takes similar approach to Dream Fields, except instead of CLIP, it uses a loss that results from distilling a 2D diffusion model. Based on probability density distillation, the loss minimizes KL divergence between a family of Gaussian distributions with common means based on the forward diffusion process and the score functions learned by the pre-trained diffusion model. By optimizing differentiable picture parameterizations, the resulting score distillation sampling (SDS) approach allows sampling. DreamFusion creates high-fidelity coherent 3D objects and scenarios for a variety of user-provided text prompts by combining SDS with a NeRF variation specifically designed for this 3D generating task. (See Fig. 1).

Natural language modeling can enable new methods of conceiving and making the world around us. Despite the notable recent improvements in text-to-image production, as discussed earlier the lack of large-scale coupled text and model data means that text-to-model creation remains a difficult task. There is a straightforward but efficient procedure for a zero-shot text-to-model creation that gets around this data shortage. A pre-trained image-text dataset like CLIP and an unlabeled model dataset are both requirements for CLIP-Forge, which is built on a two-level training task [2]. This approach has the benefit of developing several models for a given text

Fig. 1 DreamFusion using a pre-trained text-to-image diffusion model to generate realistic 3D models from text prompts

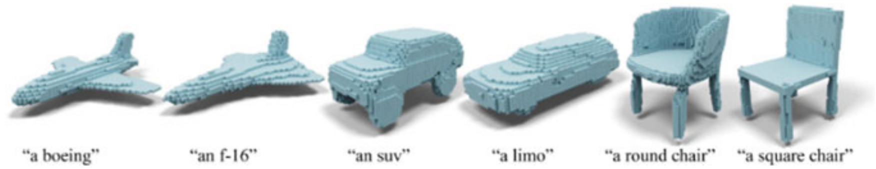


Fig. 2 CLIP-Forge generating zero-shot text-shape generation from simple categorical text prompts

without the need for costly inference time optimization (see Fig. 2). It presents thorough comparative evaluations to better comprehend its behavior and demonstrates promising zero-shot generalization of the CLIP-Forge model on both a qualitative and quantitative level.

Text-to-model generation models are a crucial enabler for new intelligent tools in design, production, as well as computer graphics and games. Text and image modalities are now connected in a significant way. Recently, DALLE and the related pre-trained image-text embedding model CLIP showed encouraging outcomes for the text-to-image generation problem. On tasks other than those for which the model was trained, they prominently displayed zero-shot generalization. The issue of developing models is more fundamental than creating images since images are projections and renderings of the fundamentally 3D physical world. Therefore, one would wonder if the 2D success can be translated into the 3D universe. This turns out to be a difficult problem. Contrary to the image generation from text case, where paired data is commonplace, it is impossible to compile huge, matched datasets of texts and models.

The main contributions of this paper are as follows:

- It introduces a novel technique called CLIP-Forge that creates 3D models straight from the text without the need for linked text-model labeling.
- It has an accurate creation process requiring no inference time optimization and produces multiple models for text.
- It provides an extensive evaluation of our method in various zero-shot generation settings qualitatively and quantitatively.

The metric for evaluation of the models is perceptual evaluation in the case CLIP-Forge. It is determined whether adding greater detail in the text affects the resulting model. In 70.83% of the image pairs, the human evaluators recognized the model created by the detailed query, demonstrating that this method can make use of attribute and internal category information in a recognizable manner. They note that models generated by attribute triggers were simpler to recognize than those generated by subcategory triggers. Given the quality of generation, one explanation for this result is that many of the subcategories are tougher to discern, whereas the attribute-augmented triggers give a precise description of how the item should appear.

3.2 Enhanced Edge Detection Techniques to Achieve Pixel-Level Accuracy

Even some of the newest end-to-end deep learning (DL) algorithms still struggle to achieve high pixel-level accuracy, making edge preservation in noisy environments a difficult issue. The improved Canny edge detector (CED) introduced in this study uses the same input parameter set as the traditional Canny but produces an edge map with better-connected edges and less noise [3]. The enhanced Canny uses stochastic resonance (SR) guided threshold maneuvering and window mapping, which is one of the most popular edge detection operators. The SR-based analysis guides the actions that need to be taken to improve the functionality of the traditional CED. In addition, it provides a brand-new method for effective edge detection, a special technique for efficiently extracting edge content and combining it with other channels, as well as a framework for dealing with the effects of noise that is random. The proposed remedy is a modular patch-based structure, making it simple to include in future algorithm modifications.

Saliency detection is about salient object detection or eye fixation prediction, with the aim of extracting the majorly dominant objects or predicting eye detectable locations corresponding to informative regions in an image. In recent years, success of fully convolutional network (FCN) in the field of computer vision, deep learning methods are preferred alternative to salient object detection. Tu et al. [4] proposed method to combine global context information and local structure information with various resolution details and integrate edge priors into different resolution features for locating the boundary of the salient object precisely than ever before [5]. Their research states that embedding an edge prior knowledge into hierarchical feature maps helps preserve better boundaries of salient captured objects.

3.3 Size-Invariant 3D Generation from a Single 2D Rock Image

2D imaging can produce inexpensive and quick data compared to the relatively expensive capture of 3D images. A challenging non-deterministic inverse task, however, is the reconstruction of a 3D image from a single 2D image [6–8].

Specifically for Rocks

Numerous deep learning-based algorithms are developed in the past, but majority of them lack the ability to generalize the shapes and qualities of various rocks, are time-consuming, and can only produce small images (3003 voxels cube). They suggested a size-invariant multi-step 3D creation procedure from a single 2D image. Since it is intended to provide statistically representative pore structures in addition to meeting the big size limitation (> 10,003 voxels cube), the suggested method addresses several significant issues in the development of 3D images. We can go over the limitations of GAN approaches in terms of scalability, stability, and complexity by combining these several generative techniques.

They used rocks with various physical features, sizes, and resolutions to train the suggested hypothesis. It developed a set of large-scale three-dimensional rock photos and compared with actual three-dimensional photographs using physical characteristics to validate their methods (porosity, permeability, and Euler characteristic). Hence, the objective of the work is to develop an automatic workflow for 3D generation from a single 2D rock image [9].

As a result, during training, the models should be solid and reliable so that a continuous training pipeline can be used to regularly integrate new samples. GAN-based techniques are constrained by this requirement since they are unstable while training with altering datasets [10].

4 Applications of 3D Models

The applications of this proposed project lie in the field of extended reality. Extended reality includes immersive areas—augmented reality and virtual reality. It makes the process of building a 3D model efficient by saving time and reducing resource usage (hardware resources such as graphic processing units and memory).

Implementable in initial stages of design for manufacturing companies such as automobiles (BMW, Ford, etc.) and aerospace industries.

5 Conclusion, Challenges, and Future Work

The best illustration of how technology brings ideas closer to reality is how 3D modeling has transformed the presentation of architectural plans. The most significant change that has occurred is the effect 3D modeling is on how architectural services are presented. 3D models are now being used on job sites and in project presentations by the architecture and construction industries. Machines and computers go through lengthy procedures to create 3D models. To better illustrate and comprehend the work at hand, 3D modeling has been employed in numerous architecture presentations.

A growing number of augmented reality applications have begun to appear in education, industry, etc., in recent years because of the development and application of AR technology. These apps display information using 3D models, such as car models to display automotive composition information and human skeleton models to display information about the human body [11]. Augmented reality (AR) systems, a brand-new interactive technique, can give people a direct-viewing sensation and more detailed information, which improves the efficiency of information comprehension. The ability of 3D technology to facilitate experiential teaching and learning in many fields where it would otherwise be difficult or impossible is one of its most notable qualities [12, 13]. The inaccessible and unseen can both become visible and accessible thanks to 3D technology.

References

1. Poole B, Jain A, Barron JT, Mildenhall B (2022) DreamFusion: text-to-3D using 2D diffusion. In: CVPR 2022. https://doi.org/10.48550/arXiv.2209.14988
2. Sanghi A, Chu H, Lambourne JG, Wang Y, Cheng C-Y, Fumero M, Malekshan KR (2022) CLIP-forge: towards zero-shot text-to-shape generation. In: CVPR 2022. https://doi.org/10.48550/arXiv.2110.02624
3. Dhillon D, Chouhan R. Enhanced edge detection using SR-guided threshold maneuvering and window mapping: handling broken edges and noisy structures in Canny edges. IEEE Access 10. https://doi.org/10.1109/ACCESS.2022.3145428
4. Tu Z, Ma Y, Li C, Tang J, Luo B. Edge-guided non-local fully convolutional network for salient object detection. IEEE Trans Circ Syst Video Technol 31(2)
5. Mukunoki M, Yasuda K, Asada N (2005) 3D model generation from image sequences using global geometric constraint. In: Advances in visual computing, pp 470–477
6. Akman A, Sahillioğlu Y, Sezgin TM (2022) Deep generation of 3D articulated models and animations from 2D stick figures. Comput Graphics 109(2022):65–74
7. Püschel H, Sauerbier M, Eisenbeiss H. A 3d model of castle Landenberg (ch) from combined photogrammetric processing of terrestrial and UAV-based images. Int Arch Photogramm Remote Sens Spatial Inf Sci XXXVII(B6b)
8. Tingdahl D, Van Gool L (2011) A public system for image based 3D model generation. MIRAGE 2011, LNCS 6930:262–273
9. Phan J, Ruspini L, Kiss G, Lindseth F (2022) Size-invariant 3D generation from a single 2D rock image. J Petroleum Sci Eng 215(Part B):110648
10. Fukatsu Y, Aono M (2021) 3D mesh generation by introducing extended attentive normalization. In: 2021 8th International conference on advanced informatics: concepts, theory, and applications (ICAICTA)
11. Choy CB, DaDanieu, Gwak J, Chen K, Savarese S (2016) 3d-r2n2: a unified approach for single and multi-view 3d object reconstruction. In: European conference on computer vision. Springer, pp 628–644
12. Wang N, Zhang Y, Li Z, Fu Y, Yu H, Liu W, Xue X, Jiang Y-G (2021) Pixel2Mesh: 3D mesh model generation via image guided deformation. IEEE Trans Patt Anal Mach Intell 43(10)
13. Pang M, He L, Xiong F, Yang X, He Z, Han X (2020) Developing an image-based 3D model editing method. IEEE Access 8
14. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S (2020) An image is worth 16x16 words: transformers for image recognition at scale. In: CVPR 2021. arXiv:2010.11929

Care A Pet—Android Pet Care Application

Kattula Chandrika, Chandrapatla Dedeepya, Gopisetti PushpaLatha, and Guttikonda Kranthi Kumar

Abstract These days, having a pet is a fun addition to everyone's life. Keeping pets has become a habit, and caring for them has become part of our human culture. To take care of your pets when you are out on a vacation, we proposed a project called "CARE A PET". The developed Android application will show the nearby pet care centers so that they can book slots according to their availability by chat. In case of an emergency time, we can call the pet specialist by taking an appointment by booking a vet on-call service using this app. The findings were divided into two sections: developing a mobile application for user advice and evaluating the application's performance for research objectives. This project also has commercial potential, compared to other methods of caring for and monitoring pets through the use of pet resources, especially in an urban way of life.

Keywords Android application · Pet caring · Pets specialist · Slot booking

1 Introduction

A domestic animal is a pet. In 2018, the estimated population of pet dogs in India was around 19.5 million. By the end of 2023, the population is predicted to approach 31 million. Each year, some 600,000 pets are adopted. Pets help their owners stay

K. Chandrika · C. Dedeepya · G. PushpaLatha (✉) · G. K. Kumar
Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India
e-mail: gopisettipushpalatha7@gmail.com

K. Chandrika
e-mail: chandrikakattula29@gmail.com

C. Dedeepya
e-mail: dedeepyachandrapatla2002@gmail.com

G. K. Kumar
e-mail: kranthi@vrsiddhartha.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_34

physically and emotionally fit. Pets can provide companionship to lonely people, mentally stressed, or elderly adults who do not have enough social interaction with others. The growing number of pets, pet owners' attitudes about pets, and the impact of fast-changing technologies all contribute to the rising need for pet care. Petting is not only profitable these days, but it has also become a new trend. Many people are concerned about how to care for their pets and where to keep them when they leave the city for a few days, and sometimes individuals are unable to farm even though they are interested in pets for several reasons, so they can farm a pet for a limited time. Having a pet may be difficult and exhausting. As a result, this software offers a simple and minimal way of taxing to deal with this scenario.

Smart mobile devices can be used to advise and direct what care to take and provide a simple solution to eliminate the problem pet owners encounter when they wish to travel out of town for a few days where they will store their pets and how to take care of themselves. As a result, this app is used to solve a problem. In this application, people who want to provide pets talk to a care center and leave pets in a care center when they leave the station. This application is an Android application, and therefore, it can be used on mobile phones at any time and from any location.

2 Literature Survey

Yasasvi Yallam and Ruchera Vaidya et al. [1] demonstrate that Happy Paws uses a database to record information about various dog and cat breeds, making it easy for consumers to contact us and purchase a pet of their choice. With only one click, pet owners may purchase pet food and accessories. Users can book a veterinarian on-call or grooming services by contacting the information provided. Advantages: Users can contact these providers with ease. Many features are available. Because most dog and cat owners are unaware of canine blood donation programs and animal blood banks, the Happy Paws application has taken the initiative to save a life. Disadvantages: There is no mapping of nearby pet care centers. There is no online consultation with veterinary doctors when needed.

Saswadkar et al. [2] exemplify the difficulty of developing an Android pet care application that would allow users to request care for their pets and find interested caregivers. The owner of the animal will receive comprehensive information about his or her pet, including details about the creature's food and hybrids, if any interested parties answer the request. The user's location will be shown before requests and feedback. The owner will give those who want compensation. Advantages: It is beneficial to pet owners. It is also a user-friendly program. Increases pet-related awareness. Disadvantage: There is no guarantee that our pet will be safe and will not be monitored.

Liyanage et al. [3] demonstrate that the goal of this research article is to examine the characteristics and uses of contemporary smart Pet Care technology. Secondary data were used in this study. A systematic approach was used to collect data from research published in Smart Pet Care Applications. Advantages: CCTV service.

Remote feeding and automatic defecation. Disadvantages: There is no friendly pet care. No doctor's services are available.

Tang et al. [4] demonstrate the Pet Tracker system enables pet owners to enter and monitor the activities of their animals inside, such as in a home or apartment. The motors in our system right now have natural sensor boards installed. One of these components is utilized to keep an eye on a pet's whereabouts, activities, and surroundings. Advantages: Pet tracking. This app is made for pet owners who wish to keep tabs on their animals, while they are away from the house. Disadvantages: When the pet is in distress, there is no one to assist it. There are no grooming services available.

Soleh et al. [5] demonstrate this application serves as a marketing tool for all pet stores in the Tangerang region that provide pet services such as grooming, food, and shopping. The Opets application can assist pet owners in finding a pet store nearby. Advantages: Buying pets. Pet grooming. Chatbox allows you to easily communicate with a pet store manager. Disadvantages: This app does not provide notifications when necessary. No veterinary services are available for pets.

Shih et al. [6] demonstrate this project presented the communication interface between a pet and its owner. In addition to the camera, heart rate and temperature sensors are implemented in a wearable device on the pet. Because the wearable device is linked to Wi-Fi, the owner can monitor the pet's situation. Advantages: This device can be loaded with modules to provide feedback to the pet. The pet can be monitored by the owner. Disadvantage: No one is available to care for the pet in the event of an emergency, and the owner cannot always supervise the pet in the house.

Kumar et al. [7] demonstrate the pet mobile app can be used to keep pets safe from disease. In an emergency, we can contact a veterinarian by scheduling an appointment online with this app. This program can assist you in locating the nearest pet hospital in your region. In this study, issues concerning pet health are addressed using mobile phones. Advantages: It saves both time and the life of the pet. Can look up doctors in a specific area and find specific hospitals. Disadvantages: Limited to relatively narrow problems. In some cases, a communication gap may exist between the customer and the doctor.

3 Proposed System

This section discusses the proposed system's architecture and methodology.

3.1 Architecture

The suggested model for the system of Care A Pet is depicted in Fig. 1. The proposed system specifies in detail the process of what would occur in the application from start to finish in the indicated way.

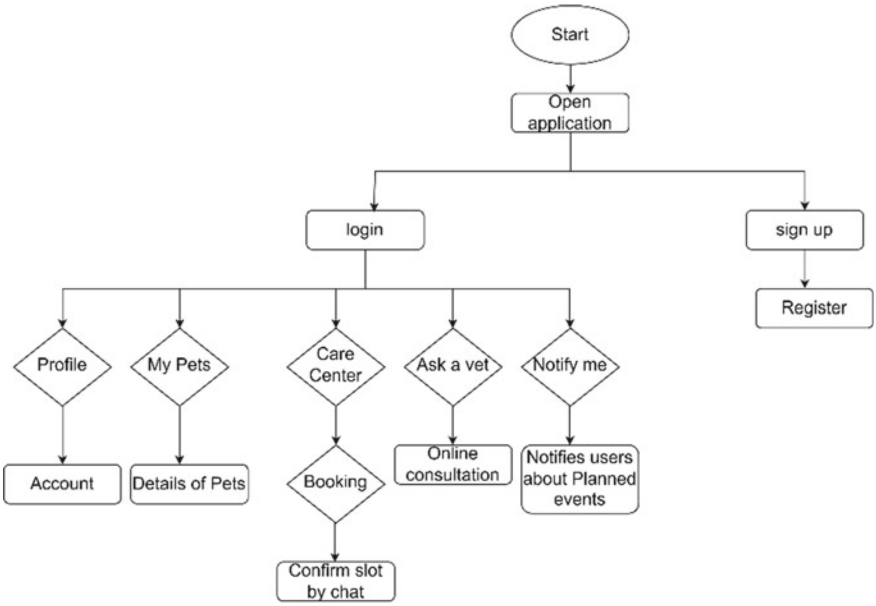


Fig. 1 Proposed system model for Care A Pet

3.2 Methodology

Care A Pet application uses a database to store descriptions about various dogs and cats of all breeds, making it easier for users to care for their pets while traveling by contacting us. Pet owners can easily book a slot by chatting with the care center and receiving notifications about their upcoming activities with a single click. Users can get veterinary and grooming services for their pets. Users can book a vet on-call or a grooming service by contacting the contact information provided.

This app is intended for pet owners who want to keep track of their pet's activity while they are away from home. This program gives friendly pet care to pet hostels, guaranteeing that their pet is in good hands. The developed Android application will display nearby pet care centers, allowing them to book slots based on their availability via chat. When necessary, online consultations with veterinary doctors are available. Through Care A Pet, pet owners can easily contact pet care centers. This application can notify pet owners of upcoming events such as scheduled vet visits, and remote communication between pet owners and pet care facilities is simple.

All of the above services are available to users through a single application. All of the following services are available simply by registering with the app, and all user information and requests are saved safely and securely in the firebase database utilizing firebase authentication and a real-time database.

3.3 Algorithm

Open the mobile application. If a new user registers into the app.
Else Login

Step 1: If the user wants to book a slot, then go to the care center module and book a slot by Whatsapp chat conservation
Step 2: If the user wants to consult a doctor online, then go to the vet details module and consult the vet by call
Step 3: To see the details of the user go to the account details module
Step 4: To see the details of the pets go to the pet details module to add new details of the pet or use the existing details
Step 5: If the user wants to get notified for predefined events, then go to the notification module for setting their remainders for notifying.

4 Results and Analysis

This section describes the system's results and output. Login page and sign up page and what other functionalities the application provides (Figs. 2, 3, 4 and 5).

5 Conclusion

Finally, the core components of the Care A Pet application are pet services, slot booking, event notifications, and a pet care center. Pet While traveling, this application provides pet owners with extensive knowledge about their pets. In this paper, we outline the design of a mobile operating system based on the Android operating system that would provide individuals with a simple means of finding information about how to take care of dogs when they are not at home. It saves us time by making it easier to get appointments online. It is extremely beneficial to society as well as the life and health of pets. However, for future tests, we hope to add new features and improve the application using some advanced techniques (such as extending the project to other care centers and chatbots for communication).

Fig. 2 Login page

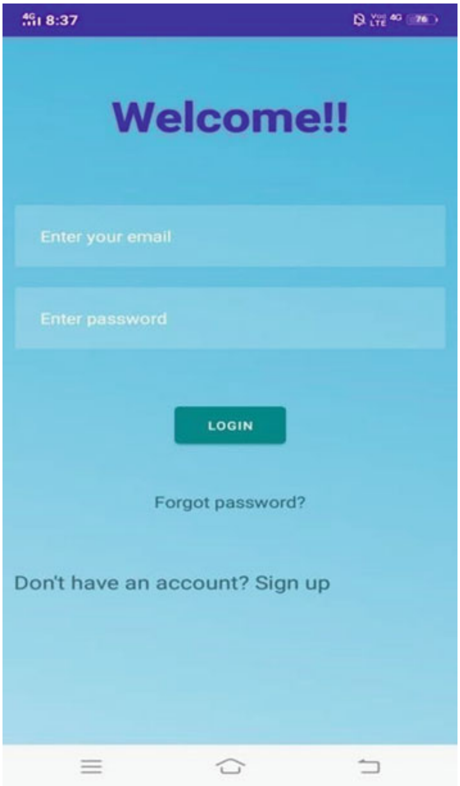
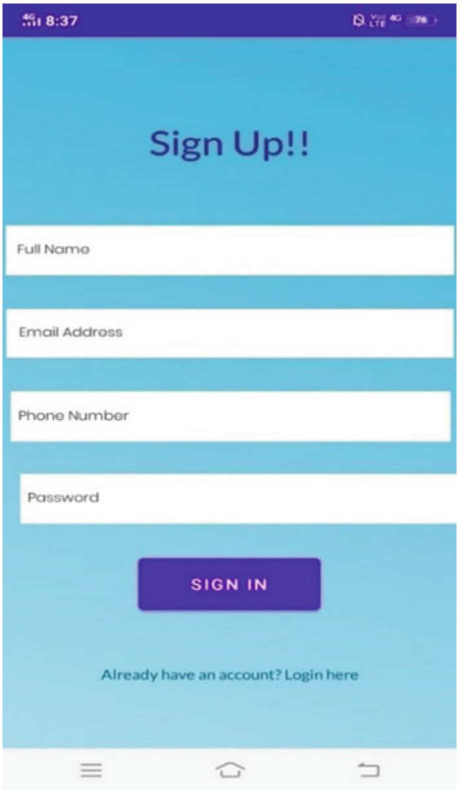


Fig. 3 Sign up page



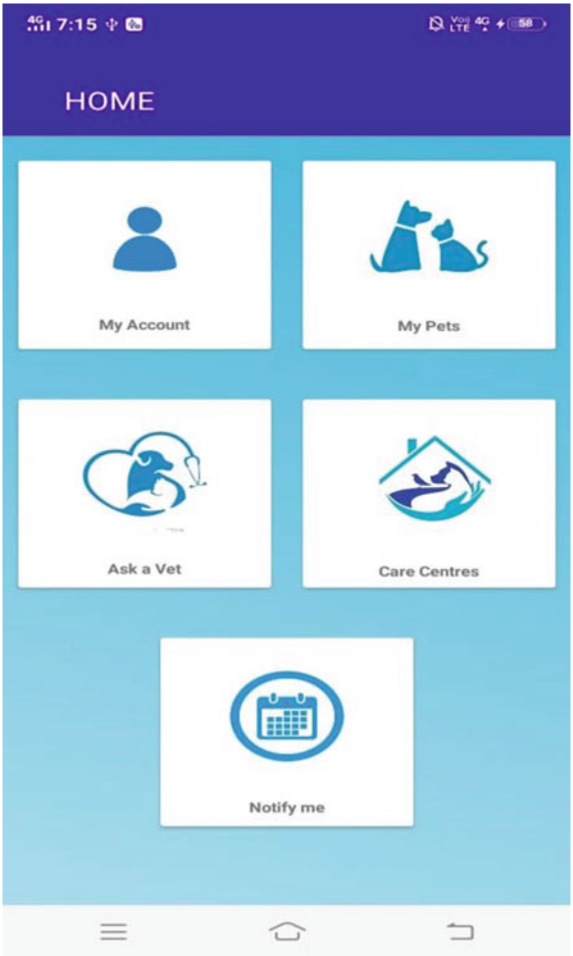


Fig. 4 Home page

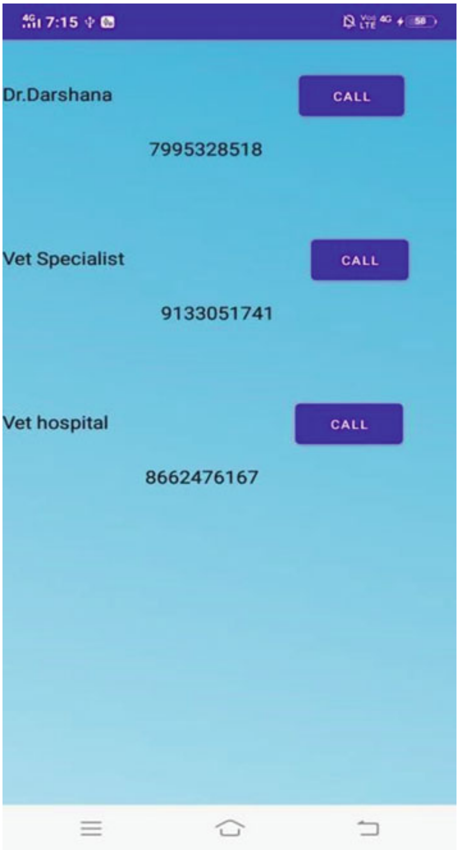


Fig. 5 Ask a vet

References

1. Yasasvi Yallam R, Ruchera Vaidya R (2020) Happy paws—android pet application. Int Res J Eng Technol (IRJET) 7(6). e-ISSN: 2395-0056, p-ISSN: 2395-0072. www.irjet.net
2. Saswadkar V, Paygude V, Dudhe P, Garad P, Gaikwad R (2018) Pet care system based on android application. Int J Trend Scient Res Dev (IJTSRD) 2(4):1488–1491. ISSN: 2456-6470. https://www.ijtsrd.com/papers/ijtsrd14415.pdf
3. Liyanage WLSV, Wedasinghe N (2021) Implementation of smart pet care applications in an IoT based environment. Published in KDUIRC international research conference 2021, ID 433,35-it0055@kdu.ac.lk
4. Tang Z, Hile H, Bajracharya S, Jurdak R (2005) PetTracker—PetTracking system using motes. Published in ResearchGate and article: January 2005, publication at: https://www.researchgate.net/publication/237368760
5. Soleh O, Wuryani R, Farizi R (2017) OPet's is petshop mobile application to meet all the needs of pets (day-care, shopping, and grooming): development and business. In: 2017 2nd International conferences on information technology, information systems and electrical engineering (ICITISEE), pp 141–146https://doi.org/10.1109/ICITISEE.2017.8285483

6. Shih Y-S, Samani H, Yang CY (2016) Internet of Things for human—pet interaction. In: 2016 International conference on system science and engineering (ICSSE), 2016, pp 1–4. https://doi.org/10.1109/ICSSE.2016.7551607
7. Kumar P, Madhu, Kumar J, Sathish C (2017) Health experts for pets using mobile apps. In: 2017 International conference on algorithms, methodology, models, and applications in emerging technologies (ICAMMAET), pp 1–2. https://doi.org/10.1109/ICAMMAET.2017.8186740

Structure Learning of Bayesian Network from the Data

Naveen Kumar Bhimagavni and T. Adilakshmi

Abstract Learning the structure of Bayesian network is a two-step process, one is parameter learning, and other is finding the best structure among search space using uncertain and incomplete data. Structure learning is the most important and complex task (NP hard problem) in estimation theory. However, existing techniques such as K3 algorithm require topological order of the nodes or a constraint on the maximum number of parents for a node, and also, most of them are generating all possible graphs even for small number of random variables and consume large amount of space and time complexity (Buntine in IEEE Trans Knowl Data Eng 8:195–210 [1]; Chickering in Learning from data. Springer, pp 121–130 [2]) to verify each of the structure. In this work, we propose an algorithm that generates comparatively small number of graphs as a heuristic search technique, and Bayesian score is calculated for each candidate structure to find the best network structure. The proposed algorithm consumes comparatively less time complexity to discover network structure using the data.

Keywords Bayesian network · Covid 19 dataset · Bayesian score · Subset graph algorithm

1 Introduction

Bayesian network is a compact representation of joined probability distribution and referred as directed acyclic graph (DAG) that consists of vertices and edges. Each vertex denotes a random variable, and edge from (A → B) represents the conditional dependency between random variables A and B. BN is a compact representation of joint probability distribution of all random variables [3].

N. K. Bhimagavni
University College of Engineering, Osmania University, Hyderabad, India

T. Adilakshmi (✉)
Vasavi College of Engineering, Osmania University, Hyderabad, India
e-mail: t_adilakshmi@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_35

1.1 Structure Learning

Bayesian network can be constructed manually by domain experts using prior structure knowledge especially in medical domain. The structure can then be refined manually or by automatically using techniques such as refinement algorithm and ExpertBayes [4]. However, manual construction may not be feasible for all domains. Another approach is to use available information. Structure of BN can also be estimated directly from the data. There are two approaches for the structure learning: score-based approach and constraint-based approach [5].

Constraint-Based Approach

The constraint-based case employs the independence test [6] to identify a set of edge constraints [7] for the graph and then finds the best DAG that satisfies the constraints [8].

Score-Based Approach

The score-based approach first defines a criterion to evaluate how well the Bayesian network fits the data and then searches over the space of DAGs [9] for a structure with maximal score [10].

Most of the existing techniques [1] follow the below steps to find the best structure. First, generate all possible graphs (which consumes large time complexity [2]) and then apply any scoring function for each candidate structure; the structure with the highest Bayesian score is considered as the best Bayesian network structure. However, the scoring-based approach generates exponent number of candidate graphs.

2 Proposed Algorithm

In this work, a new heuristic search technique is proposed that generates optimal number of graphs, and then, scoring function is applied on each of the graph to find the best structure.

Step 1: Generate optimal number of candidate structure.
Step 2: Calculate Bayesian score for each candidate structure using sufficient statistics of the input data.
Step 3: The structure with the highest Bayesian score is the best Bayesian network structure.

Algorithm 1: Proposed algorithm

```
Data:
   Input data for covid 19
Result:
   Bayesian Network structure
      1. begin
```

```
    2. Read Input data for covid 19;
    3. Best_Score = 0;
    4. Best_structure = 0;
    5. Calculate Sufficient statistics (S) for the input data;
   6. Call G = GenerateOptimalGraphs(n); // n is number of random
variables in the Input data
    7. for each graph g_i in G
begin
     8. Calculate Bayesian Score(BS(g_i))for each graph g_i using
Sufficient statistics (S);
    9. if BS(g_i) > Best_Score then
  Best_Score = BS(g_i);
  Best_structure = g_i
end
    10. end
```

3 Calculate Bayesian Score for Each Structure

Bayesian score [3] can be used as scoring function for all possible graphs [11, 12]. It is the function which takes one graph as an input and computes the Bayesian score for that structure [13, 14].

$$P\left(\frac{G}{D}\right) = \frac{P\left(\frac{D}{G}\right)P(G)}{P(D)} \tag{1}$$

$P(D \mid G)$ is the marginal likelihood of the graph structure G. $P(G)$ is the prior probability of the graph structure and can be considered as uniform distribution G in the candidate structure space. $P(D)$ is the probability of the data which acts as a normalizing constants the score function can be written as below; it primarily depends on the marginal likelihood $P(D \mid G)$

$$\mathrm{ScoreB}(G : D) = \log P\left(\frac{D}{G}\right) \tag{2}$$

3.1 Data Preprocessing

The data was collected from source kaggle [15] for the pandemic disease Covid-19 that consists of 35 million records; the dataset contains 27 features with missing values. These values are handled using Python package. Test data contains missing

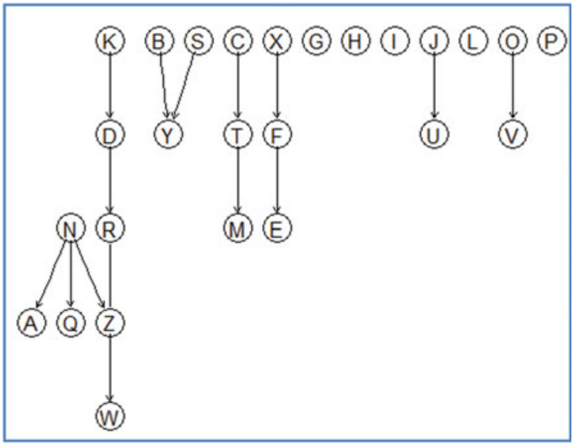


Fig. 1 Best Bayesian network structure for covid_ 19 dataset created by proposed algorithm

values for mixed data types (numeric or strings), and SimpleImputer from the Python package sklearn.impute can be used.

Ex: imp_mean = SimpleImputer (missing_values=np.nan, strategy="most_ frequent")

When strategy is set to "most_frequent", it replaces missing values using the most frequent value along each column. It can be used with strings or numeric data (mixed data types) (Fig. 1).

4 Goodness Fit of the Model

Proposed heuristic algorithm can be compared with the existing structure learning algorithms such as hill-climbing learnt by bnlearn package. Bayesian network models can be compared with the following metrics.

1. score comparison
2. hold-out cross-validation
3. precision.

1. Score Comparison

The Bayesian scoring function is used to compute the score of the Bayesian network with 27 nodes learnt using bnlearn package (hill-climbing approach)

$$\text{score}(\text{bnlearn\_BN}, 26 \text{ nodes}) = 3240.513$$

Score of the Bayesian network with 20 nodes learnt using heuristic algorithm

$$\text{score}(\text{heuristic\_BN}, 26 \text{ nodes}) = 3576.921$$

2. Hold-Out Cross-Validation

Cross-validation is a standard method used to estimate a model's goodness of fit. In k-fold cross-validation, data is randomly divided into k subsets. For each subset k, a model is evaluated on k having been trained on $(k - 1)$ subsets (Fig. 2; Tables 1 and 2).

3. Precision Analysis

Precision: It computes the proportion of positive identifications was actually correct out of total predicted positives. 3×3 confusion matrix created for the node sore throat which has three levels, namely low, moderate and high, diagonal elements represent the true positives along that column, and the remaining values in the same column represent false positives. Precision values have been computed with various

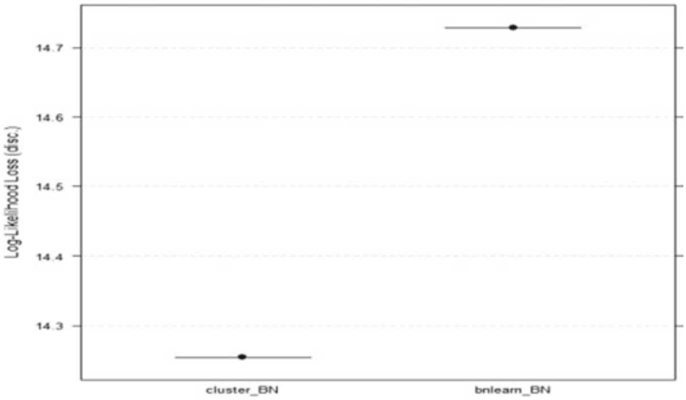


Fig. 2 Expected loss comparison graph for bnlearn_BN and heuristic_BN

Table 1 Expected loss for the Bayesian network created by bnlearn package

Number of folds in cross-validation (k)	$K = 10$
Number of observations that are to be sampled for the test subsample	$m = 50{,}000$
Loss function	Log-likelihood loss
Expected loss	14.80372

Table 2 Expected loss for the Bayesian network created by heuristic algorithm

Number of folds in cross-validation (k)	$K = 10$
Number of observations that are to be sampled for the test subsample	$m = 20{,}000$
Loss function	Log-likelihood loss
Expected loss	14.65423

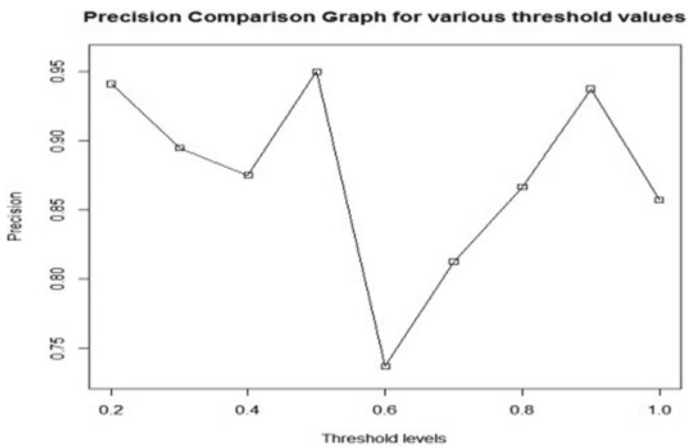


Fig. 3 Precision values computed for different threshold values (0.2–1.0)

Table 3 3 × 3 confusion matrix created for the node sore throat

Observed values	Predicted values		
	Low	Moderate	High
Low	0	1	1
Moderate	0	8	0
High	0	0	4

threshold values ranging from 0.2 to 1.0 for the below confusion matrices from precision comparison graph shown in Fig. 3 (Table 3).

5 Conclusion

The proposed algorithm generates less number of candidate structures, thereby consumes comparatively less time complexity and finds the best structure in optimal time. Bayesian score is used as a scoring function to compute the score of the each candidate structure. The goodness fit of the best Bayesian network structure is measured in terms of Bayesian score, expected loss and precision values. It is observed that score of the best Bayesian network structure has improved by 10%. At the same time, expected loss is reduced by 1.0098 percentages.

However, present work can be extended for large number of nodes in the graph, and Bayesian score can be calculated for Bayesian networks having both discrete and continuous random variables.

References

1. Buntine W (1996) A guide to the literature on learning probabilistic networks from data. IEEE Trans Knowl Data Eng 8(2):195–210
2. Chickering DM (1996) Learning Bayesian networks is NP-complete. In: Learning from data. Springer, pp 121–130
3. Koller D, Friedman N (2009) Probabilistic graphical models: principles and techniques. MIT Press
4. Beretta S, Castelli M, Gonçalves I, Henriques R, Ramazzotti D. Learning the structure of Bayesian networks: a quantitative assessment of the effect of different algorithmic schemes
5. Bernstein D, Saeed B, Squires C, Uhler C (2020, June) Ordering-based causal structure learning in the presence of latent variables. In: International conference on artificial intelligence and statistics, pp 4098–4108. PMLR
6. Chen Y, Tian J (2014, June) Finding the k-best equivalence classes of Bayesian network structures for model averaging. Proc AAAI Conf Artif Intell 28(1)
7. Chen EYJ, Shen Y, Choi A, Darwiche A (2016) Learning Bayesian networks with ancestral constraints. Adv Neural Inf Process Syst 29:2325–2333
8. Cheng J, Greiner R, Kelly J, Bell D, Liu W (2002) Learning Bayesian networks from data: an information-theory based approach. Artif Intell 137(1–2):43–90
9. Chickering D (2002) Learning equivalence classes of Bayesian-network structures. J Mach Learn Res 2:445–498
10. Chickering DM, Meek C (2002, August) Finding optimal Bayesian networks. In: Proceedings of the eighteenth conference on uncertainty in artificial intelligence, pp 94–102
11. Chobtham K, Constantinou AC (2020) Bayesian network structure learning with causal effects in the presence of latent variables. In: Proceedings of the 10th international conference on probabilistic graphical models, in proceedings of machine learning research, vol 138, pp 101–112
12. Claassen T, Mooij JM, Heskes T (2013, August) Learning sparse causal models is not NP-hard. In: Proceedings of the twenty-ninth conference on uncertainty in artificial intelligence, pp 172–181
13. Colombo D, Maathuis MH (2014) Order-independent constraint-based causal structure learning. J Mach Learn Res 15(1):3741–3782
14. Colombo D, Maathuis MH, Kalisch M, Richardson TS (2012) Learning high-dimensional directed acyclic graphs with latent and selection variables. Ann Stat:294–321
15. Kaggle data source for Covid19. https://www.kaggle.com/iamhungundji/covid19-symptoms-checker

Self-adaptive Bald Eagle Search for Energy Efficient Load Balancing in Cloud System

Shilpa B. Kodli and Sujata Terdal

Abstract By distributing incoming traffic among several servers, networks, or other resources, load balancing helps organizations manage workload demands while enhancing performance and preventing service interruptions. This research work follows 3 modules for load balancing including: workload submission, central manager (CM), as well as migration. Cloud users submit the tasks during the submission of workloads. Migration as well as allocation of resources takes place in the second stage. Employing Pearson Correlation Coefficient (PCC) technique, a unique safety constraint is used. A new Self Adaptive Bald Eagle Search (SA-BES) optimization algorithm is introduced for choosing the VM and for allocating the resources. Moreover, Global Agent (GA) as well as correlation-based VM placement applies during the load balancing. Lastly, actual migration is performed by the VM from the schedule gathered from the GA.

Keywords Work load · Global agent (GA) · Self adaptive bald eagle search (SA-BES) · Virtual machine (VM) · Pearson correlation coefficient (PCC)

1 Introduction

Cloud computing resources offers internet services to users with high speed. Cloud offers storage as well as hosting services on Internet, which seems to be a practical growth of automation. This demand was virtualized, location-independent, and paid using the cost approach in an effort to achieve high performance and optimal resource usage [1]. However, there are some challenges, such as resource utilization, QoS (Quality of Service), security, load balancing and fault tolerance [2, 3]. In recent years, it has become incredibly popular. Cloud computing services can be used on a personal or business level. The distribution of workloads over numerous virtual machines (VMs) in order to offer the necessary IT resources and improve

S. B. Kodli (✉) · S. Terdal
Department of CSE, P. D. A College of Engineering, Kalaburagi, Karnataka, India
e-mail: shilpakodli@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_36

overall system performance is one of the main issues with the cloud [4, 5]. Load balancing admits to spreading resources among several cloud nodes with the goals of maximizing service utilization, enhancing system performance, and conserving energy.

The name "load balancing" is used in terms of factors including data processing time as well as VM response time, which have no effect on data centres when requests for time changes are made in the cost calculation of the algorithm [6–8], which is demonstrated in considerations of load balancing that is effective. Additionally, load balancing speeds up cloud computing by lowering execution time, waiting time, response time, and throughput. It is also the most used simplicity technique and is sufficient to be efficient in any case. In order to move the loaded process time between jobs and processors, the load balance has indeed been employed during optimization [9, 10]. By utilizing the autonomy function, VM will adjust to changes in their surroundings and avoid being over or underloaded. This will help keep the system stable and limit requests moving across VMs, lowering network overheads in the process. This research's primary goal is:

- Introducing fivefold multi-objective constraints namely: DI, makespan, completion time, RU, as well as EC for the optimal VMs allocation.
- Introduces a SA-BES algorithm for optimal allocation of VM.

Remaining section of this work is organized as: review on existing topic is given in Sect. 2. Section 3 describes the VM consolidation technique along with SA-BES algorithm introduction for VM migration as well as optimal VM allocation through PCC. Sections 4 and 5 describes the obtained results as well as conclusion.

2 Literature Review

In 2021, Annie Poornima Princess and Radhamani [11] created a hybridized HHO (Harries Hawks Optimization) and PIO (Pigeon Inspired Resource Consumption Optimization) load balancing method to provide an effective load balance which improvises key factors, such as a fast allocation of resources. By identifying the overload and underload situations of VMs (Virtual Machines), HHO can allocate tasks, and a PIO-based method can speed up its response time.

In 2020, Neelima and Rama Mohan Reddy [12] created a novel load balancing job scheduling technique in the cloud to address the NP-hard load balancing optimization issue. The ADA (Adaptive Dragonfly algorithm) is produced from the combination of the Dragonfly as well as Firefly methods. In order to improve performance, a multi-objective function focused on three parameters completion time, operational cost, and load was constructed.

In 2020, Negi et al. [13] developed a CMODLB (Clustering-based Multiple Objective Dynamic Load Balancing) method for cloud load balancing in the current study. First, previously developed ANN-LB (Artificial Neural Network-based Dynamic Load balancing) method is introduced for VMs clustering into the overloaded and under loaded VMs via BOEK-means (Bayesian Optimization-based Enhanced K-means) technique. Secondly, scheduling the user tasks for the under loading VMs for enhancing the resource usage and load balancing. The multi-objective TOPSIS-PSO (Technique of Order Preference by Similarity to Ideal Solution with Particle Swarm Optimization) method, which uses a number of cloud conditions, aids in task scheduling.

In 2020, Ebadifard and Babamir [14] created a dynamic scheduling strategy that considers the request compatibility with the selected VM as relying on resource vector rather than processing for the actual distribution of VM requests. Several techniques developed for problem of transmission overheads is ignored. Various experiments are conducted here for addressing this issue via automated Load Balancing algorithm. The focus of the available research on cloud computing task scheduling is primarily on CPU-bound requests.

In 2021, Velpula and Pamula [15] introduced an EBGO technique which tackle the drawbacks of current load balancing techniques in regards of particular important parameters including: response time, weighted overall cost, energy usage, procession period etc. This method tries to achieve complete optimal parameter values, successfully balancing server demand in data centres to produce optimal resource consumption. This method is appropriate for medical care systems because a lot of different kinds of systems and data were being exchanged between them.

3 Suggested VM Consolidation Method in Cloud

We developed a SLA aware dynamic VM consolidation technique depending on load balancing. Developed technique incorporates 4 stages: (1) Initial VMs allocation on unoccupied hosts utilizing SA-BES technique. (2) Host overload as well as under load prediction technique via PCC. (3) VM selection technique employing imbalance degree with MAD [9]. (4) Placement of VM via PCC for optimizing the VM allocation.

3.1 VM Consolidation Architecture

This research follows 3 modules which includes: workload submission, CM, and migration. In the first module, cloud users hand over a variety of tasks for VMs to complete. Using the data gathered by local agents, the CM is responsible for managing live migration, resource allocation, and total system resources in the following module. The best migration strategy is offered by a global agent after the first VM allocation depending on SA-BES optimization, the host detection component's evaluation of the existing state of each host's load and identification of the set of overloaded and under-loaded hosts, and the PC's use of this data to optimize the placement of the VMs have all been completed. The global agent uses the placement controller and VM allocation balancer in two phases to accomplish the goal. Global agent first asks load balancer SA-BES to keep the system balanced while initially allocating VMs to various hosts. The host is monitored by the local agent in the final module, and the local agent transmits monitored data to the global agent, who then gives the best possible migration strategy.

3.2 Introduction to SA-BES Algorithm for Optimal VM Allocation

We developed a SA-BES method in this section for optimal VMs allocation on unoccupied PMs.

VMs Allocation in the PMs. Through the development of SA-BES optimization, the VM is allocated while taking into account 5 PM constraints including: DI (Degree of Imbalance), VM-CT (VM-Completion Time), MS (Makespan), RU (Resource Consumption), EC (Execution Cost).

DI: The amount of load allocated amongst the VMs in relation to their implementation capabilities is measured and evaluated as given in Eq. (1), where $\mathrm{tot}_{\max}$, $\mathrm{tot}_{\min}$, $\mathrm{tot}_{\mathrm{avg}}$ represents the utmost, least as well as mean total execution time between overall VMs. The lower DI value indicates that the load is well balanced.

$$DI = \frac{tot_{\max} - tot_{\min}}{tot_{avg}} \tag{1}$$

VM-CT: It is defined as the execution period following completion of the last task, and it is calculated as shown in Eq. (2), here n, m represents the VM as well as task count.

$$\mathrm{ct}_i = \sum_{j=1}^{m} \frac{t_j.\mathrm{Len}}{\mathrm{vm}_i.\mathrm{PEN} \times \mathrm{VM}_i.\mathrm{mips}}, \quad \text{with } i \in \{1, 2 \ldots n\} \tag{2}$$

MS: It is defined as the total ct needed to complete each task and is evaluated as shown in Eq. (3), here n represents the VM total. The lower MS shows that the scheduler is providing resources with an effective and good allocation of tasks.

$$\mathrm{MS} = \max_{1 \le i \le n} \{\mathrm{ct}_i\} \tag{3}$$

RU: It is employed to evaluate resource usage. A higher percentage of resource exploitation shows that cloud providers might make lots of profit. In Eq. (4), RU is evaluated.

$$RU = \frac{\sum_{i=1}^{n} ct_i}{MS \times n} \tag{4}$$

EC: It compares the entire cost of payments made by cloud customers to cloud providers to the resources used to carry out tasks. Users' primary goal while using the cloud is to save costs while utilizing it effectively and with the least MS. In Eq. (5), EC is computed.

$$\mathrm{EC} = \sum_{i=1}^{n} \mathrm{Price}_i^{*} \mathrm{ct}_i \tag{5}$$

Objective Function and Solution Encoding. This study tries to reduce SLA violations by allocating VMs into appropriate PMs in an efficient manner while taking into account five constraints. In this case, the resource usage must be increased while the degrees of imbalance, MS, execution time and VM must be minimized. In Eq. (6), research study objective is provided where weights $W{*}_1$, $W{*}_2$, $W{*}_3$, $W{*}_4$, $W{*}_5$ were set as 0.2. As a result, if SA-BES is given 10 VMs as input, the chromosome size will also be 10.

$$\mathrm{obj} = \min\left\{ \begin{array}{l} W *_1 .\mathrm{DI} + W *_2 .\mathrm{MS} + W *_3 . \\ \mathrm{ct} + W *_4 .(1 - \mathrm{RU}) + W *_5 .\mathrm{EC} \end{array} \right\} \tag{6}$$

Proposed SA-BES model for optimal Allocation. The nature-inspired meta-heuristic optimization method SA-BES [10] imitates the bald eagles' social intelligence and hunting tactics as they look for fish.

(i) **SA-BES select stage**: Bald eagles seek and select the best hunting area within the selected search space depending on the food availability in the select step. In Eq. (7), this behavior is mathematically presented, where, ρ indicates the control parameter for variations in position which ranges from [1.5, 2] as well as Rd indicates the random number which ranges from [0, 1]. Choosing the current search area (M_{best}) by bald eagles based on the best location discovered over their previous search. M_{mean} means that these eagles have digested all of the information from the prior points.

$$M_{\mathrm{new},i} = M_{\mathrm{best}} + \rho * \mathrm{Rd}(M_{\mathrm{mean}} - M_i) \tag{7}$$

(ii) **SA-BES search stage**: Bald eagles travel in a spiral pattern within the chosen search zone during the search stage in order to speed up their hunt for prey. In

Eq. (8), swoop best position is expressed mathematically, where g indicates the parameter which ranges from [5, 15] for defining corner amongst the central point searches and F ranges from [0.5, 2] for defining the search cycle count.

$$M_{i,\text{new}} = M_i + L(i) * (M_i - M_{i+1}) + K(i) * (M_i - M_{\text{mean}}) \tag{8}$$

Here, $K(i) = \frac{\text{Kr}(i)}{\max(|\text{Kr}|)}$, $L(i) = \frac{\text{Lr}(i)}{\max(|\text{Lr}|)}$, $\text{Kr}(i) = r(i) * \sin(\theta(i))$, $\text{Lr}(i) = r(i) * \cos(\theta(i))$, $\theta(i) = g * \pi * \text{Rd}$, $r(i) = \theta(i) * F * Rd$

(iii) **SA-BES swooping stage**: During the swooping stage, bald eagles soar to their targeted prey from the best spot in the search region. Furthermore, the ideal point is where all points converge. Equation (9) provides a mathematical example of this behavior.

$$\begin{aligned} M_{i,\text{new}} = &\text{Rd} * M_{\text{best}} + K1(i) * (M_i - l1 * M_{\text{mean}}) \\ &+ L1(i) * (M_i - l2 * M_{\text{best}}) \end{aligned} \tag{9}$$

Here, $K1(i) = \frac{\text{Kr}(i)}{\max(|\text{Kr}|)}$, $L1(i) = \frac{\text{Lr}(i)}{\max(|\text{Lr}|)}$, $\text{Kr}(i) = r(i) * \sinh(\theta(i))$, $\text{Lr}(i) = r(i) * \cosh(\theta(i))$, $\theta(i) = g * \pi * \text{Rd}$, $r(i) = \theta(i)$ Where, $l1, l2 \in [1, 2]$.

According to SA-BES, probability pr is calculated in the Eq. (10). If $\text{pr} \le 0.5$, position update is done using Eq. (11) and if $\text{pr} > 0.5$, position update is done using Eq. (12). Here, Rd indicates the chaotic random number logistic map.

$$\text{pr} = -\exp\left(1 - \frac{\text{TT}}{\text{tt}_{\max}}\right)^{\frac{1}{20}} \tag{10}$$

$$\begin{aligned} M_{i,\text{new}} = &\text{Rd} * M_{\text{best}} + K1(i) * (M_i - l1 * M_{\text{mean}}) \\ &+ L1(i) * (M_i - l2 * M_{\text{best}}) + \text{Levy}(\beta) \end{aligned} \tag{11}$$

$$\begin{aligned} M_{i,\text{new}} = &\text{Rd} * M_{\text{best}} + K1(i) * (M_i - l1 * M_{\text{mean}}) \\ &+ L1(i) * (M_i - l2 * M_{\text{best}}) + \text{Brownian motion} \end{aligned} \tag{12}$$

As per SA-BES, arithmetic crossover operation is performed. The usage of arithmetic crossover is possible in the context of real-value encoding. The arithmetic crossover [16] operator linearly combines the two parent chromosomes. In an arithmetic crossover, 2 chromosomes are randomly selected for crossover, and a correlation between two of these chromosomes results in the creation of two offspring. This linear combination comes about as a result of the calculation below:

$$\begin{aligned} &\text{Ch}\,1 = u \cdot P1\,\text{gene} + (1 - u) \cdot P2\,\text{gene} \\ &\text{Ch}\,2 = u \cdot P12\,\text{gene} + (1 - u) \cdot P1\,\text{gene} \end{aligned} \tag{13}$$

Algorithm 1: **Pseudo code of SA-BES algorithm**

Initialize the point M_i for n and Evaluate the initial point fitness: $f(M_i)$
While end condition is not met
for each i point in population
Select stage position update is done using Eq. (7)
if $f(M_{\text{new}}) \prec f(M_i)$, $M_i = M_{\text{new}}$
if $f(M_{\text{new}}) \prec f(M_{\text{best}})$, $M_{\text{best}} = M_{\text{new}}$
end if
end if
end for
for each i point in population
Search stage position update is done using Eq. (8)
if $f(M_{\text{new}}) \prec f(M_i)$, $M_i = M_{\text{new}}$
if $f(M_{\text{new}}) \prec f(M_{\text{best}})$, $M_{\text{best}} = M_{\text{new}}$
end if
end if
end for
for each i point in population
Calculate probability pr as per SA-BES
if pr ≤ 0.5, Swoop stage position update is done using Eq. (11) as per SA-BES
else if pr $\succ 0.5$, Swoop stage position update is done using Eq. (12) as per SA-BES
Assign $x = x + 1$
end while

3.3 VM Migration Through PCC

VM Migration includes following steps:

(a) **PM load recognition and computation technique**

This method aims to reduce the number of migrations and SLA violations while enhancing performance of cloud computing. PCC [17] is a measure used in statistics to assess the strength of correlations between variables. The PCC is used in this study to determine how closely each VM's computational resources match those of each host (PM). In this study, three categories of computing resources are examined: memory RAM (R), bandwidth (BW) and CPU (cpu). According to available resources, Eq. (14) is used to calculate (PCC). PCC (o, p) in Eq. (14) indicates association degree between the o, p resources. Z_i^o, Z_i^p represents the equal value of the resources o, p, Z^{-p}, Z^{-o} indicates the p, o related mean value for VM_i, as well as k represents the VM total in particular host.

$$\text{PCC}(p, o) = \frac{\sum_{i=1}^{k}\left(\left((Z_i^o - Z^{-o}) \times \left(Z_i^p - Z^{-p}\right)\right)\right).}{\sqrt{\sum_{i=1}^{k}\left(Z_i^o - Z^{-o}\right)^2} \times \sqrt{\sum_{i=1}^{k}\left(Z_i^p - Z^{-p}\right)^2}} \tag{14}$$

In order to accomplish its goal, upper- and lower-thresholds are established for detecting under- and overloaded hosts, respectively. Due to the variety of computational resources, PCC is used to determine the safety constraint, as shown in Eq. (14).In Eq. (15), safety constraints (sc) computations are given. Additionally, in Eqs. (16) and (17), upper threshold (UT) as well as lower threshold (LT) were defined.

$$\mathrm{sc} = \left[\mathrm{PCC}(\mathrm{cpu}, R) * \mathrm{PCC}(\mathrm{cpu}, R)\right] + \frac{\left[\mathrm{PCC}(\mathrm{cpu}, \mathrm{BW}) * \mathrm{PCC}(\mathrm{cpu}, \mathrm{BW})\right] + \left[\mathrm{PCC}(R, \mathrm{BW}) * \mathrm{PCC}(R, \mathrm{BW})\right]}{2} \tag{15}$$

$$\mathrm{UT} = 1 - sc \times \hat{o}(p_{k+1}) \tag{16}$$

$$\mathrm{LT} = \min \mathrm{Ut} - \mathrm{sc} \times \min \mathrm{Ut} \tag{17}$$

The current host (PM) usage (ChUt) is calculated after the thresholds have been calculated, as shown in Eq. (18). It split MIPS provided by the total MIPS consumed by all of the VMs operating on the host. The method then calculates the utilization rate (Utrate) from the $ChUt$ total. In Eq. (19), computation Utrate is calculated, where Ut = Total Requested Mips/host.get Total Mips defines the host usage.

$$\mathrm{ChUt} = \mathrm{Ut} + \mathrm{Utrate} \tag{18}$$

$$\mathrm{Utrate} = \mathrm{sc} \times \mathrm{Ut} \tag{19}$$

If ChUt > UT, then assign host as overloaded host
If ChUt < LT, then assign host as under load host.

(b) **VM Selection Strategy**

Select the VM to be moved from the host if it turns out that the host is overloaded according to the data in the aforementioned section. An important step in maximizing VM migration as well as energy usage is VM selection. In order to ensure lower SLA violations and energy usage, the maximal Mean Deviation Utilization (MDU) specified in Eq. (20) is deployed along with the minimal Migration Delay (MD). According to Eq. (21) the MD is calculated using ram and bandwidth. So, the VM with the lowest MD and highest MDU is selected.

$$\mathrm{MDU} = \underset{1 \le i \le m}{\mathrm{MAD}_i} * \mathrm{VM}^{\mathrm{IB}} \tag{20}$$

$$\mathrm{MD} = \frac{\mathrm{VM}.R}{\mathrm{VM}.B} \tag{21}$$

(c) **VM Placement Scheme**

After choosing the virtual machine (VM) for migration, the new host is chosen for the placement of VM based on the host's processing power. Overall VM is first arranged in decreasing order of CPU utilization. Then, use PCC between RAM and CPU after computing the mean power use variance among hosts with and without recently arriving VM. Here, two PCC between RAM as well as CPU are estimated as given in Eqs. (22) and (23), respectively. Each VM is allotted to a host with a slight increase in mean power consumption difference (MPD). PCC^{k+1}, PCC^{k} represents the PCC with $k + 1$ VMs as well as k VMs, respectively. The VM is allocated to a host based on MPD.

$$\text{PCC}^{k} = \text{PCC}(\text{cpu}, R) \tag{22}$$

$$\text{PCC}^{k+1} = \text{PCC}(\text{cpu}, R) \tag{23}$$

4 Results and Discussion

The proposed load balancing in cloud system was simulated in CloudSim and implemented using JAVA. The dataset used is PlanetLab dataset which is assembled in [18]. The SA-BES was evaluated with the traditional methodologies, such as PRO (Poor and Rich optimization), BM (Blue Monkey), AOA (Arithmetic Optimization Algorithm), SOA (Seagull Optimization Algorithm), and BES (Bald Eagle Search), respectively. Here, when the task is set to as 2000, then the SA-BES and other schemes were assessed using two experiments (a) Number of VM = 1516 and (b) Number of Host = 800, respectively. The proposed was examined with respect to degree of imbalance, energy consumption and VM migration.

4.1 Performance Analysis on SA-BES When the Number of VM is 1516

See Table 1.

The analysis on SA-BES and traditional methodologies (PRO, BM, AOA, SOA and BES) is evaluated in terms of degree of imbalance, energy consumption and VM migration. Moreover, the assessment is performed for diverse number of hosts ranging from 800, 950, 11,000, 1250 and 1400, when the VM is fixed to as 1516 at a task count 2000. The SA-BES has accomplished with minimum degree of imbalance, energy consumption and VM migration in almost all the hosts. Therefore, the

Table 1 Performance evaluation on varied metrics using SA-BES versus existing methods for varied number of host

Task = 2000, Number of VMs = 1516						
Number of hosts	PRO	BM	AOA	SOA	BES	SA-BES
Degree of imbalance						
800	2516.4	2452.39	2206.485	1647.829	1639.664	1606.041
950	2713.8	2499.67	2483.531	2470.409	2402.286	2397.830
1100	2950.3	2883.46	2865.045	2730.395	2604.274	2460.422
1250	3181.8	3045.78	2946.592	2934.896	2851.079	2719.242
1400	3220.8	3144.12	2985.097	2941.469	2881.387	2783.455
Energy consumption (kWh)						
800	124.2277	118.218	113.6586	108.5863	104.0404	99.98064
950	148.5277	141.512	136.9847	132.3059	121.2806	114.4747
1100	180.0754	165.696	151.7069	143.4768	132.2375	120.3986
1250	203.3521	184.175	180.6693	163.1931	158.9543	139.8688
1400	227.8625	214.949	209.5465	203.9332	195.6399	181.5353
VM migration						
800	7521	7213.177	7078.847	6616.702	5943.397	4870.045
950	7511	7476.989	7445.190	7437.586	7391.362	7372.156
1100	7443	7270.392	7106.218	6993.850	6911.427	6790.955
1250	7387	7371.298	7349.683	7333.148	7327.556	7306.792
1400	7279	7168.433	7086.485	7066.995	7018.381	7005.403

extensive enhancement in those metrics makes the SA-BES more effectual at load balancing in cloud.

4.2 Performance Analysis on SA-BES When the Number of Host is 800

The degree of imbalance, energy consumption and VM migration have estimated in this section, and the appropriate outcomes of the SA-BES are compared to the PRO, BM, AOA, SOA and BES are illustrated in Table 2. Additionally, the number of VM is adjusted to 900, 1000, 1100, 1200 and 1300. The number of VM is set to as 1100, the degree of imbalance of the SA-BES is 2751.565, though the conventional methods scored higher degree of imbalance notably, PRO = 3401.655, BM = 3266.40264, AOA = 3207.99807, SOA = 3007.50277 and BES = 2989.2437. Thus, it exhibits the improvement of the SA-BES method for load balancing in cloud system with lower energy consumption, degree of imbalance and VM migration.

Table 2 Performance evaluation on varied metrics using SA-BES versus existing methods for varied number of VM

Task = 2000, Number of host = 800						
Number of VM	PRO	BM	AOA	SOA	BES	SA-BES
Degree of imbalance						
900	3907.579	3160.8929	2880.12888	2678.6345	2599.5195	2565.918
1000	3514.7	3475.00307	3271.52012	3236.4601	3208.5392	3057.246
1100	3401.655	3266.40264	3207.99807	3007.5027	2989.2437	2751.565
1200	3204.713	3014.0136	2722.43155	2683.9157	2646.7151	2472.160
1300	2923.562	2886.30181	2775.19635	2761.8513	2687.6314	2625.378
Energy consumption (kWh)						
900	194.5683	188.866354	176.603369	150.19560	127.62105	120.0396
1000	178.2213	162.508868	160.034915	151.87596	140.37711	125.5640
1100	166.4167	162.508868	160.034915	151.87596	140.37711	125.5640
1200	150.2357	144.675169	141.142651	133.78278	129.61398	117.2670
1300	143.9321	138.902306	136.183039	134.06274	127.81864	113.8848
VM migration						
900	4293	4288.427	4260.93932	4260.8664	4251.2621	4206.150
1000	4820	4807.21918	4706.85783	4603.8929	4546.1957	4365.283
1100	5357	5290.01104	5163.97205	5052.1889	4988.784	4585.987
1200	5879	5798.27111	5791.60893	5787.2206	5785.9786	5711.320
1300	6399	6214.68903	5868.00302	5546.5762	5544.7386	5297.749

5 Conclusion

Our study follows three modules including: workload submission, central manager (CM), as well as migration. Cloud users submit the tasks during the submission of workloads. Migration as well as allocation of resources was taken place in the second stage. Employing Pearson correlation coefficient (PCC) technique, a unique safety constraint is developed. A new self-adaptive bald eagle search (SA-BES) optimization algorithm is introduced for choosing the VM and for allocating the resources. Moreover, global agent (GA) as well as correlation-based VM placement applies the load balancing. Lastly, actual migration is performed by the VM from the schedule gathered from the GA.

References

1. Kaur A, Kaur B, Singh D (2019) Meta-heuristic based framework for workflow load balancing in cloud environment. Int J Inf Tecnol 11:119–125
2. Kong L, Mapetu JPB, Chen Z (2020) Heuristic load balancing based zero imbalance mechanism in cloud computing. J Grid Computing 18:123–148
3. Abdulhammed O (2022) Load balancing of IoT tasks in the cloud computing by using sparrow search algorithm. J Supercomput 78:3266–3287
4. Hung LH, Wu CH, Tsai CH, Huang HC (2021) Migration-based load balance of virtual machine servers in cloud computing by load prediction using genetic-based methods. IEEE Access 9:49760–49773
5. Shafiq DA, Jhanjhi NZ, Abdullah A, Alzain MA (2021) A load balancing algorithm for the data centres to optimize cloud computing applications. IEEE Access 9:41731–41744
6. Sohani M, Jain SC (2021) A predictive priority-based dynamic resource provisioning scheme with load balancing in heterogeneous cloud computing. IEEE Access 9:62653–62664
7. Kruekaew B, Kimpan W (2022) Multi-objective task scheduling optimization for load balancing in cloud computing environment using hybrid artificial bee colony algorithm with reinforcement learning. IEEE Access 10:17803–17818
8. Nabi S, Ibrahim M, Jimenez JM (2021) DRALBA: dynamic and resource aware load balanced scheduling approach for cloud computing. IEEE Access 9:61283–61297
9. Mapetu JPB, Kong L, Chen Z (2021) A dynamic VM consolidation approach based on load balancing using Pearson correlation in cloud computing. J Supercomput 77:5840–5881
10. Alsattar HA, Zaidan AA, Zaidan BB (2020) Novel meta-heuristic bald eagle search optimisation algorithm. ArtifIntell Rev 53:2237–2264
11. Annie Poornima Princess G, Radhamani AS (2021) A hybrid meta-heuristic for optimal load balancing in cloud computing. J Grid Computing 19(21)
12. Neelima P, Reddy ARM (2020) An efficient load balancing system using adaptive dragonfly algorithm in cloud computing. Cluster Comput 23:2891–2899
13. Negi S, Rauthan MMS, Vaisla KS (2021) CMODLB: an efficient load balancing approach in cloud computing environment. J Supercomput 77:8787–8839
14. Ebadifard F, Babamir SM (2021) Autonomic task scheduling algorithm for dynamic workloads through a load balancing technique for the cloud-computing environment. Cluster Comput 24:1075–1101
15. Velpula P, Pamula R (2022) EBGO: an optimal load balancing algorithm, a solution for existing tribulation to balance the load efficiently on cloud servers. Multimed Tools Appl 81:34653–34675
16. Furqan M, Erianto Ongko H, Ikhsan M (2017) Performance of arithmetic crossover and heuristic crossover in genetic algorithm based on alpha parameter. IOSR J Comp Eng (IOSR-JCE) 19
17. Xu J, Tang B, He H, Man H (2016) Semi supervised feature selection based on relevance and redundancy criteria. IEEE Trans Neural Netw Learn Syst 28(9):1974–1984
18. https://planetlab.cs.princeton.edu/datasets.html

Survey on Estimation of Body Mass Index Using Facial Images

Salakapuri Rakesh, Keerthi Aluvala, S K M Aqeel, and Madireddy Harika

Abstract One of the most concerning public health issues the world is currently dealing with is obesity. Body mass index (BMI) is used to measure one's obesity and is useful for evaluating an individual's health. BMI is calculated by dividing a person's weight by the square of their height. Calculating BMI from just a facial image has been a challenge in the field of computer vision, and researchers have used various methodologies. In this paper, various methods used by different researchers have been described and summarized according to the research objectives. Researchers have used two major methods for extracting features from facial images, i.e., extracting geometric features and deep features. According to our research, region-aware global average pooling together with deep features were recommended, and training of dataset based on gender and race was recommended which improves accuracy and gives the best results for estimation of BMI from facial images.

Keywords Facial features · Obesity · Body Mass Index · Geometric features · Deep features

S. Rakesh · K. Aluvala (✉) · S. K. M. Aqeel · M. Harika
Department of IT, CBIT, Hyderabad, India
e-mail: ugs19134_it.keerthi@cbit.ac.in

S. Rakesh
e-mail: srakesh_it@cbit.ac.in

S. K. M. Aqeel
e-mail: ugs19151_it.aqeel@cbit.ac.in

M. Harika
e-mail: ugs19130_it.harika@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_37

1 Introduction

Over the years, overweight and obesity rates among adults have increased constantly [1]. Obesity can cause various health concerns like diabetes, cancer, heart disease, Alzheimer, depression, sleep apnea, artery disease, stroke, and bone and joint problems [2, 3]. There have been various reasons for this increasing rate of obesity. The advancement of technology has had a great contribution to it. As technology advances, day-to-day activities are getting simpler, hence reducing human activity which is leading to a sedentary lifestyle. With the widespread usage of smart phones, studies have shown a relationship between screen media exposure and increased risks of obesity.

Body mass index is the measure of an individual's body fat. It was given by Lambert Adolphe Jacques Quetelet [4, 5]. It was firstly observed as a disease by the World Health Organization (WHO) in 1948 [6]. Though BMI cannot differentiate between the fat and muscle associated with weight, it is still the most useful index for measurement of obesity [7]. The World Health Organization [8] has defined BMI categories as underweight, normal weight, overweight, moderately obese, severely obese, and very severely obese [9]. BMI is calculated by dividing a person's weight in kilograms by the square of their height in meters. But according to the English system, the formula for BMI is weight in pounds divided by square of height in inches and multiplied by 703 (conversion factor) [10]. BMI can also be known by referring to the Center for Disease Control and Prevention (CDC) table with given weight and height [11].

$$\mathrm{BMI} = \frac{\mathrm{Weight(kg)}}{\{\mathrm{Height(m)}\}^2}$$

Human faces present a number of cues such as gender, age, expression, and personality attributes which are used in recognition of identities [12], face expression detection [13], face detection for security, face tracking [14, 15], for identification of visual attributes such as race, age and gender [16] and for deepfake detection [17, 18].

Researchers have now come up with the relation of facial features with body weight and body mass index. Lots of research has been carried out in which BMI is estimated from facial images. With the growing health concerns, the need of an easy self-health monitoring tool is vital. Enabling users to have an easy mechanism by which they can view their BMI using facial images and self-monitor themselves is the need of the hour.

2 Prior Work

In this section, we will discuss the earlier research on the estimation of BMI from facial images using various deep learning and machine learning architectures and models.

Wen and Guo [19] proposed the first study, to our best knowledge, on the prediction of BMI using facial images. The active shape model (ASM) was used by the authors to extract facial landmarks from facial images, and after further processing, they extracted seven geometry-based features.

Yousaf et al. [20] extracted deep features, employed face semantic segmentation, and obtained pixel-wise localization of various facial regions. The authors have used FaceNet and VGG-Face for deep feature extraction.

Siddiqui et al. [21] developed a custom end-to-end CNN network for the prediction of BMI using facial images. The authors compared and evaluated the performance of ResNet-50, VGG-19, DenseNet-121, lightCNN-29, lightCNN-29, and MobileNet-V2 for extracting facial features. The authors used ridge regression (RR) and support vector regression (SVR) to predict BMI from extracted features. They showed that ResNet and DenseNet models performed better when used with ridge regression, and that pre-trained models performed marginally better than the end-to-end CNN model.

Sidhpura et al. [22] used tensor processing units (TPU) and created Tensorflow records for the images of the dataset for faster processing. They have used Inception-v3, VGG-Face, VGG-19, and Xception models for feature extraction.

Pham et al. [23] used three different architectures of ResNet which are ResNet-152, ResNet-101, and ResNet-50 to extract feature vector from a facial image. They proposed a new regression module to predict BMI from a feature vector.

Jiang in [24] compared both deep learning-based features and geometric features. For geometric features, they analyzed two methods that are psychology-inspired geometric features (PIGF) and pointer feature (PF). For feature extraction using deep learning-based methods, they have used the models, namely VGG-Face, ArcFace, Centerloss, and LightCNN. Features were extracted, and SVR has been used as a regression model. The authors used two datasets, one which they have collected and named the FIW-BMI and a pre-existing Morph-II dataset. They concluded that the performance of deep learning models is better than geometric-based models on large datasets. Among the deep models, Arcface and VGG-Face gave more reliable results.

Min Jiang [25] have used a two-stage learning framework for BMI prediction using facial images. The two stages are feature learning and estimator learning. The first stage consists of learning of BMI features. The second stage consists of learning a BMI estimator based on label distribution by an optimized method in which the features and assigned labels were projected to a new domain which resulted in maximized correlation between them. The authors tested the results on FIW-BMI, VIP_attribute, and Morph II datasets. VIP_attribute and Morph-II are existing datasets, whereas the FIW-BMI dataset was collected by authors which contains 7930 images. They have concluded that the two-stage learning framework reduces the error significantly and performs better than two deep learning, regression-based and two label distribution methods.

3 Datasets

The datasets used by various researchers have been described in this section.

Morph-II Dataset [26]

The original Morph-II dataset contains 55,352 images in which there are 42,722 samples belonging to black, 10,655 white, 57 Indian, 160 Asian, 1,753 Hispanic, and 5 belonging to other categories. Out of the total samples, only 37,626 have their respective height and weight data. The dataset is not publicly available.

Bollywood Dataset [27]

The Bollywood dataset consists of 237 images of Bollywood celebrities. It contains images of 22 celebrities, hence containing multiple images of each celebrity. All dataset consists of only Indian images. This dataset is available publicly at Github [27].

VIP Attribute Dataset [25]

This dataset consists of facial images with their weight, height, and respective BMI. It contains 1026 samples with an equal division of both male and female samples, i.e., 513 samples of male images and 513 samples of female images. The dataset is collected by the authors of [25] and is not publicly available.

Illinois DOC Labeled Faces Dataset [28]

This dataset is available publicly at [28]. Illinois Department of Corrections is the source of the dataset. It consists of front and side view facial images of 68,149 prisoners along with their gender, date of birth, height, and weight. It consists of some images that did not have their respective weight and height data and some corrupt images. Upon trimming such images, the dataset is left with around 56,200 male samples and 3649 female samples.

Arrest Records Dataset [29]

This dataset consists of prisoners with their names, weight, height, BMI, and other information. It has 1543 images in which 1243 subjects are male and 300 subjects are female. Mean value of BMI is 26.41. The standard deviation of BMI is 5.27. The range of the BMI values in this dataset is [15, 56].

Visual BMI Dataset [30]

This dataset consists of 4206 subjects with gender and their BMI values. In this dataset, 7 subjects fall under the range of underweight, i.e., having BMI values in the range of [16, 18.5]. Six hundred and eighty subjects fall under the category of normal weight, i.e., [18.5, 25]. One thousand one hundred and fifty-one subjects are overweight category, i.e., [25, 30]. Nine hundred and forty-one subjects are moderately obese, i.e., [30, 35]. Six hundred and eighty-one subjects are severely obese, i.e., [25, 40]. Seven hundred and forty six are very severely obese (BMI > 40).

4 Methodology

In this section, we will discuss the methodologies done by the researchers.

The authors in [19] have used AdaBoost classifier for face detection, and they have normalized detected faces which results in the fiducial point detection being more robustness and accuracy. Active shape model (ASM) was used for detection of the fiducial points in the facial images. Principal component analysis (PCA) of the active shape model was used for the position of facial components such as nose, eyes, and lips**.** The detected fiducial points have been used as input for calculating geometric or ratio features. They have used seven geometric features which are perimeter-to-area ratio (PAR), face width-to-lower face height ratio (FW/FH), cheekbone to jaw width (CJWR), mean of eyebrow height (MEH), lower face-to-face height ratio, width-to-upper facial height ratio (WHR WHR), and eye size (ES). They have used three regression methods which are least squares estimation (LSE), Gaussian process (GP), and the support vector regression (SVR) [5] and compared their performance for prediction of BMI. They divided the MORPH-II dataset into two sets, Set1 and Set2 for training and testing, and they have shown (i) the Pearson's correlation coefficient r between facial features and BMI on Set1 and Set2 (ii) Pearson's correlation coefficient r between the facial features and the BMI across different age groups (iii) Pearson's correlation coefficient r between facial features and BMI in gender–ethnicity groups.

Authors of [20] have utilized facial regions to predict BMI. They obtained various face regions using face semantic segmentation. They have obtained improved feature vectors by pooling the convolutional feature maps according to different semantic face regions. The approach contains three major components. A face detector is applied to localize the face, and deep features are extracted from facial images. Facial feature extraction models named FaceNet and VGG-Face are employed. Then, the face is localized into different regions using a modified version of BiSeNet that generates a separate binary mask for each region rather than a combined mask for all regions. The face regions segmented are ear, eyes, hair, eyebrow, neck, lips, nose, skin, and background.

Face region recognition features are then obtained by performing element-wise multiplication of the feature map using FaceNet or VGG-Face using the mask obtained from the semantic segmentation. Finally, the region-aware feature vector is employed with a regression module that gives out the predicted BMI. Figure 1 shows the approach used by the authors.

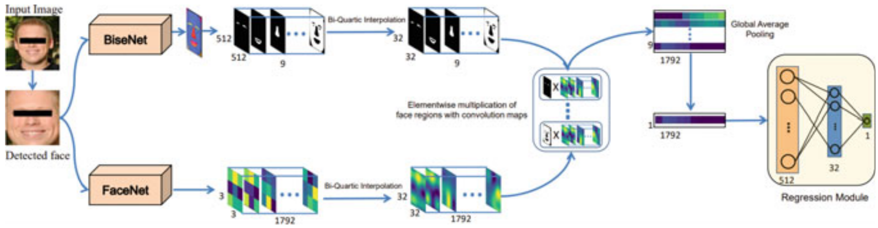


Fig. 1 Approach to prediction of BMI using facial semantic segmentation

Authors of [22] performed preprocessing over the StyleGan FFHQ dataset. They have used DLIB 68 landmark detection to vertically align faces and to blur surroundings to focus on faces. The authors used tensor processing units (TPUs) for the creation of Tensorflow records for the images in the dataset for faster processing. Every record of the Tensorflow records was resized to 256 $\times$ 256 $\times$ 3 and preprocessed for transfer learning. It contains 1024 images. They have used Inception-v3, VGG-Face, VGG-19, and Xception models for feature extraction. The same fully connected layers were used at the end of all pre-trained models. They have given the output of a pre-trained model to fully connected layers after performing Global Average Pooling. To prevent overfitting, they added a 50% dropout layer to the model. Activation function named Gaussian error linear unit (GELU) has been used. It combines the features of the activation function RELU, Zoneout, and Dropout. For fully connected layers, higher learning rate was used. For pre-trained model final layers, much lower learning rate was used which helps in extracting more features from the images. Adam optimizers were used by authors which had a decreasing learning rate in the deep layers, MultiOptimizer which was provided by TensorFlow Addons were used for implementation.

The authors in [23] have used three different architectures of residual networks for feature extraction that are ResNet-50, ResNet-101, and ResNet-152. Their research showed that ResNet's pre-trained models give better results than deep learning models, and hence, they have used ResNet models. They have proposed a new regression model which includes several blocks, and they can be selected from one to six. Each block consists of one or more sub-blocks, and each sub-block has four layers. After each dense layer, they added a group of other layers such as activation, batch normalization, and dropout to each sub-block. The dense layer size can be selected from different values. They have trained their framework in an end-end approach. Figure 2 shows the proposed regression module.

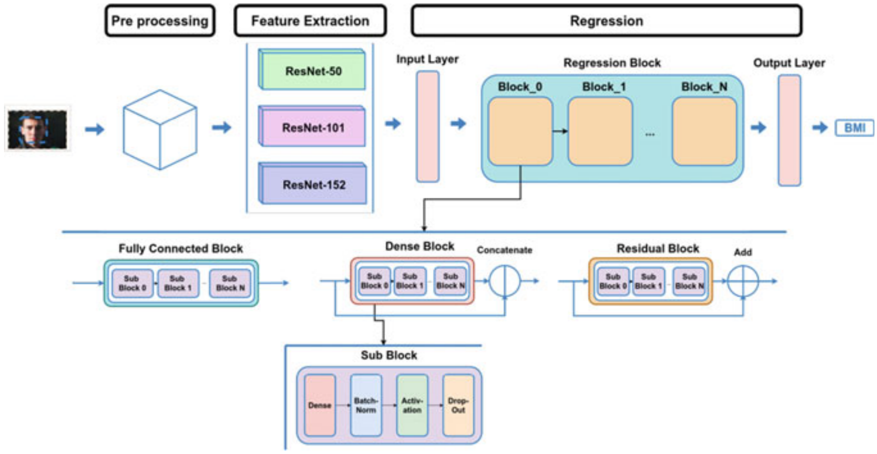


Fig. 2 Regression module

5 Result Analysis

Figure 3 shows the overall MAEs of BMI values predicted using SVR, GP, and LSE in [19]. From the results, they concluded that among three regression models SVR performed best by giving the best results for both the sets of data. They have done separated training and mixed training for gender, ethnicity, and for different age groups. MAE results of separated training of ethnicity groups using SVR are [2.78, 4.18], while results of mixed training of ethnicity groups using SVR are [2.65, 4.29]. The SVR has the lowest errors for underweight and normal, while the GP has the least errors for overweight and obese categories.

The region-aware global average pooling (Reg-GAP) approach proposed in [20] was evaluated using the following three datasets: VisualBMI, Bollywood, and VIP attribute datasets. The results showed that the Reg-GAP method gave better results than GAP and other regression methods on all three datasets. Results on the Bollywood dataset gave a Pearson's correlation of 0.55 using VGG-19 and 0.32 using VGG-Face models. An MAE of 5.03 using FaceNet and 4.99 using VGG-Face was obtained on the VisualBMI dataset. Reg-GAP outperformed GAP for all the classes (normal, overweight, obese, severely obese, and very severely obese) for VisualBMI dataset. An MAE of 1.73 was obtained on the VIP attribute dataset.

Figure 4 shows the results of the paper [22]. MAE-Overall score for Illinois DOC dataset was in the range of [2.82, 3.63]. The range of MAE-Male is [2.79, 3.61], and the range of MAE-Female is [3.54, 4.04]. Xception model among all the models used gave the best score on all three cases (i.e., MAE-Overall, MAE-Male, and MAE-Female). After the Xception model, Inception-v3 gave closer results. For all models, the MAE-Male values are slightly lower than the overall MAE, and the MAE-Female values are higher than the overall MAE. This is because the dataset is out of balance. MAE-overall for VIP_Attribute dataset was in the range of [3.10, 3.91]. The range

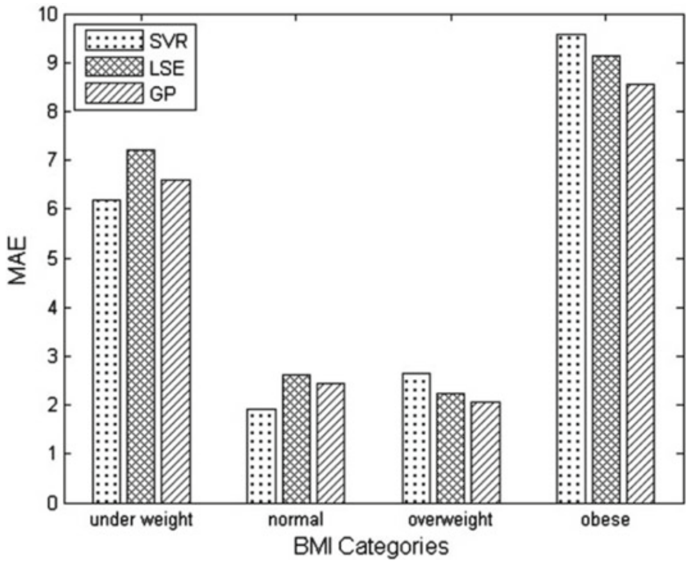


Fig. 3 MAEs visualization of SVR, GP, and LSE in various BMI categories

of MAE values of Male obtained was [2.78, 4.16], and the range of MAE values of female obtained was [2.68, 5.03]. Inception-v3 model gave the best scores among all the used models on all 3 cases. After Inception-v3, VGG-19 performed the best results on all 3 cases. The range of MAE-Overall score for arrest records dataset was [3.73, 3.93].

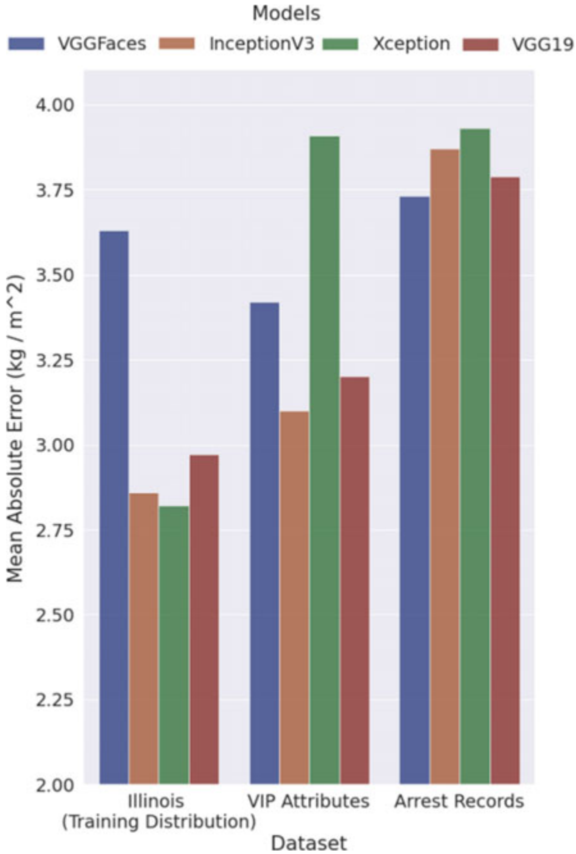
The range of MAE for male was [3.35, 3.75], and the range of MAE for female was [3.99, 5.11]. VGG-Faces gave better results for MAE-Overall. The range of MAE-Overall score for arrest records dataset was [3.73, 3.93]. The range of MAE for male was [3.35, 3.75], and the range of MAE for female was [3.99, 5.11]. VGG-Faces gave better results for MAE-Overall.

Authors of [23] have tested the results of the regression model proposed by them. The range of MAE obtained from their results is [1.83, 2.58]. They performed the model proposed by them on a small dataset which was built by themselves, and their model gave a better result on 3 classes (i.e., underweight, healthy, and overweight). They demonstrated that their proposed model produces more stable results.

6 Summary

See Table 1.

Fig. 4 Overall MAE on all datasets



7 Conclusion

After our research, we conclude that for feature extraction, region-aware global average pooling gave the best results which was done by Yousaf et al. [20]. Regression model proposed by Pham et al. [23] gave better results than other regression models. So, we conclude that the framework with feature extraction as region-aware global average pooling method with the regression model proposed by Pham et al. [23] would give better results. The authors in [20] have used BiseNet with Reg-GAP. Models which are better than BiseNet can also be used with Reg-GAP and can multiply the results with other deep CNN pre-trained models.

Table 1 Summary of the prior studies on the prediction of BMI based on facial images in terms of machine/deep learning models, dataset used, and the obtained results on BMI prediction

S. No.	Reference	Datasets	Feature type (Extraction model)	Classification/ Regression module	Results
	Wen and Guo [19]	Morph-II	PIGF (ASM)	SVR	Overall MAE: [3.05–4.29]
	Yousaf [20]	Bollywood dataset, VIP-Attributes, VisualBMI	Deep features (FaceNet, VGG-Face)	Three layer (512, 256, 1) regression module	Overall MAE: [0.32–5.03]
	Siddiqui et al. [21]	VisualBMI, VIP-Attributes, Bollywood dataset	Deep features (VGG-19, DenseNet-121, ResNet-50, MobileNet-V2, and LightCNN-29)	SVR, RR	Overall MAE: [1.04, 6.48]
	Sidhpura et al. [22]	Illinois DOC labeled faces dataset, VIP-Attributes, Arrest records dataset	Deep features (Xception, VGG-Face, Inception-v3, and VGG-19)	End-to-end	Overall MAE: [2.82, 3.93]
	Pham et al. [23]	VIP-Attributes	Deep features (ResNet-50, ResNet-101, and ResNet-152)	New regression model proposed by the authors	Overall MAE: [1.83,2.58]

(continued)

Table 1 (continued)

S. No.	Reference	Datasets	Feature type (Extraction model)	Classification/ Regression module	Results
	Jiang et al. [24]	FIW-BMI, MORPH-II	PIGF, PF, PIGF + PF (Openface) Deep features (Centerloss, Arcface, LightCNN-29, and VGG-Face)	SVR	Overall MAE Morph-II: [2.30 ± 0.03 to 3.77 ± 0.08] FIW-BMI: [3.15 ± 0.07 to 4.26 ± 0.08]
	Jiang et al. [25]	FIW-BMI, MORPH-II, VIP-attributes	PIGF, PF, PIGF + PF (Openface) Deep features (Centerloss)	SVR, PLS, GPR, LD-CCA, PCA-SVR, CCA, LD-PLS	Best Overall MAE Morph-II: (LD-CCA)—2.42 VIP attribute: (LD-CCA)—2.23

Abbreviation *ASM*, Active Shape Model; *PF*, Pointer Features; *PIGF*, Psychology Inspired Geometric Features; *PCA*, Principal Component Analysis; *RR*, Ridge Regression; *SVR*, Support Vector Regression; *CCA*, Canonical Correlation Analysis; *LD*, Label Distribution; *MAE*, Mean Absolute Error; *PLS*, Partial Least Square Analysis; *GPR*, Gaussian Process Regression

References

1. Omar HA (2009) Obesity and adolescence: a public health concern. Nova Biomedical Books
2. National Heart, Lung, Blood Institute, National Institute of Diabetes, and Kidney Diseases (US) (1998) Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. Number 98. National Heart, Lung, and Blood Institute
3. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, Hu FB, Hubbard VS, Jakicic JM, Kushner RF et al (2014) 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the american college of cardiology/american heart association task force on practice guidelines and the obesity society. J Am College Cardiol 63(25 Part B):2985–3023
4. Nuttall FQ (2015) Body mass index: obesity, BMI, and health: a critical review. Nutr Today 50(3):117–128
5. Daniels SR (2009) The use of BMI in the clinical setting. Pediatrics 124:S35
6. Philip W, James T (2008) Who recognition of the global obesity epidemic. Int J Obesity 32(7):S120–S126
7. World Health Organization (2000) Obesity: preventing and managing the global epidemic. World Health Organization, Geneva, Switzerland
8. WHO (1995) Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee, World Health Organization Tech Report Ser 854:1–452
9. Muralidhara D (2007) Body mass index and its adequacy in capturing body fat. Thai J Physiol Sci 20:97–100
10. Dietz W (2015) Using the BMI for age growth charts. Centers Dis Control Prev:29
11. U.S. Department of Health & Human Service, "Body mass index table," National Heart, Lung, and Blood Institute, https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi_tbl.htm. Last accessed 15 Nov 2022
12. Albiero V, Zhang K, Bowyer KW (2020) How does gender balance in training data affect face recognition accuracy? In: 2020 IEEE international joint conference on biometrics (IJCB). IEEE, pp 1–10

13. Zebrowitz LA (2006) Finally, faces find favor. Soc Cogn 24(5):657–701
14. Khairunnisaa A, Basah SN, Yazid H, Basri HII, Yaacob S, Chin LC. Facial-paralysis diagnostic system based on 3D reconstruction. AIP Conf Proc 1660(1):070026
15. Levine J, Ray M, Jensen M (1998) Relation between chubby cheeks and visceral fat. N Engl J Med 339(26):1946–1947
16. Krishnan A, Almadan A, Rattani A (2020) Understanding fairness of gender classification algorithms across gender-race groups. In: 2020 19th IEEE international conference on machine learning and applications (ICMLA), pp 1028–1035
17. Nadimpalli AV, Rattani A (2022) On improving cross-dataset generalization of deepfake detectors. arXiv preprint arXiv:2204.04285
18. Ramachandran S, Nadimpalli AV, Rattani A (2021) An experimental evaluation on deepfake detection using deep face recognition. CoRR, abs/2110.01640
19. Wen L, Guo G (2013) A computational approach to body mass index prediction from face images. Image Vis Comput 31(5):392–400
20. Yousaf N, Hussein S, Sultani W (2021) Estimation of bmi from facial images using semantic segmentation based region-aware pooling. Comp Biol Med:104392
21. Siddiqui H, Rattani A, Kisku DR, Dean T (2020) AI-based BMI inference from facial images: an application to weight monitoring. arXiv preprint arXiv:2010.07442
22. Sidhpura J, Veerkhare R, Shah P, Dholay S (2022) Face to BMI: a deep learning based approach for computing BMI from face. In: 2022 International conference on innovative trends in information technology (ICITIIT), pp 1–6. https://doi.org/10.1109/ICITIIT54346.2022.9744191
23. Pham Q-T, Luu A-T, Tran T-H (2021) BMI estimation from facial images using residual regression model. In: 2021 International conference on advanced technologies for communications (ATC), pp 140–145. https://doi.org/10.1109/ATC52653.2021.9598340
24. Jiang M, Shang Y, Guo G (2019) On visual BMI analysis from facial images. Image Vis Comput 89:183–196
25. Jiang M, Guo G, Guowang M (2020) Visual BMI estimation from face images using a label distribution based method. Comput Vis Image Underst 197:102985
26. Ricanek K, Tesafaye T (2006) Morph: a longitudinal image database of normal adult age-progression. In: IEEE conference on AFGR, pp 341–345
27. Bollywood dataset. https://github.com/abhaymise/Face-to-height-weight-BMI-estimation. Last accessed 20 Nov 2022
28. Fisher DJ (2019, December) Illinois DOC labeled faces dataset, Version 1, https://www.kaggle.com/davidjfisher/illinois-doc-labeled-faces-dataset. Last accessed 20 Nov 2022
29. Domestic Violence Database. National violent offender & domestic violence registry, domesticviolencedatabase.net. [Online]. Available: https://domesticviolencedatabase.net/registry_ia/?co=Polk&abc=A#abc. Last accessed 18 Nov 2022
30. Kocabey E, Camurcu M, Ofli F, Aytar Y, Marin J, Tor-ralba A, Weber I (2017) Face-to-BMI: using computer vision to infer body mass index on social media. arXiv:1703.03156

Detecting Security Breaches on Smart Contracts Through Techniques and Tools a Brief Review: Applications and Challenges

Adla Padma and R. Mangayarkarasi

Abstract In the recent decade, blockchain became a promising platform in the perception of security in almost all sectors, especially in the financial sector its contribution is phenomenal. Blockchain-enabled applications doesn't need third-party intervention while ensuring transparency since all the operations pertaining to use cases are well designed through smart contracts. A smart contract is designed to facilitate, verify, and execute the agreement between multiple untrusted parties, with tamper-resistance. Though the applications run on blockchain platforms are tamper-proof, the intruder may find unnoticed vulnerabilities in a smart contract to jeopardize the system. This paper highlights the importance of deterministic smart contracts along with the anticipated attacks. Then, a discussion on the detection of attacks using various techniques and tools is presented. Finally, research gaps identified are manifested to guide the researcher to precede further on smart contract security breaches.

Keywords Smart contract · Vulnerabilities · Use cases · Tools

1 Introduction

Blockchain technology became more popular in recent decades due to its wide range of applicability in vast domains. Bitcoin was the first cryptocurrency used by blockchain ledger technology. Bitcoin has been used to store immutable records

Supported by organization VIT-VELLORE

A. Padma · R. Mangayarkarasi (✉)
School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India
e-mail: rmangayarkarasi@vit.ac.in

A. Padma
e-mail: adla.padma2020@vitstudent.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_38

called blocks, which are connected to the previous blocks using cryptography techniques. Every transaction is signed with public-key cryptography (PKC), and details of the transactions are stored in blocks linked through cryptography algorithms. A ledger acts as a platform to store all the transactions. The special node is called a miner in the network. Every miner collects the pending transactions from the block, once the size of the collected transactions reached a predefined block size then-new a block will be created [1], in the blockchain every block has a hash of the previous block, excluding the genesis block which eliminates the merkle tree, miners, and previous hash. Data immutability is ensured through a hashing algorithm, an attempt to change the data in the block will change its associated hash value. If nodes are increased in the blockchain automatically authentication time is high due to the blockchain's linear data structure [2]. Another perspective of blockchain that ensures the consensus mechanism, existing BC implements mainly either Proof of Work (PoW) or Proof of Stake (PoS). The earlier mechanism demands more computational resources to mine the next block. Whereas in PoS, miners with a high stake will find a turn to mine the next block [3].

Smart contract (SC) plays a vital role in executing the operations on blockchain networks. Since blockchain adapted to various domains, hence, the role of SC and its potential contribution is also important. The first rule for designing an SC is, it should be deterministic. Though much attention is given to designing the SC, still it prone to security breaches. The other aspects anticipate through the smart contract, is higher performance. In 2016, Decentralized Autonomous Organization (DAO) attack took place and incurred a million Ether loss, just by manipulating SC, through re-entrancy vulnerability [4].

The rest of the paper is organized as follows. Section 2 discusses the smart contract vulnerabilities along with the existing countermeasures. Section 3, enlightens the readers with the various testing tool meant to detect the vulnerabilities in the smart contract which may lead to security breaches are furnished. Section 4 demonstrates some of the use cases enabled through smart contracts along with the challenges presented. Section 5 concludes the work.

2 Smart Contract Vulnerabilities

This section portrays the smart contract vulnerabilities demonstrated in various venues. The smart contract term was coined by Nick Szabo in 1990 but became popular in 2015 because of Ethereum Turing's complete scripting language. SC permits self-executing lines of code between two trusted third parties, and trusted transactions and agreements are carried among faceless parties without including central authority and legal system. Smart contracts are a set of rules coded using a programming language. Twenty-three languages are driving the blockchain-based application using smart contracts. Solidity is a leading language meant for developing smart contracts. Solidity has necessary programming constructs, through which smart contracts can be designed with logic schema, properties, and ledgers. The contract owner deploys and gets SC address, and then the user can invoke with available functions by sending transactions that cannot be modified by others. Implementing the operations through smart contracts brings benefits like the elimination of the third party, cost reduction, immutability, forge resistance, and transparency. A smart contract is managed with two types of accounts, namely contract accounts and externally owned accounts. The transactions are stored in the block as per the arrival time. Upon successful completion, the particular transaction will be removed by the state machine. The built-in smart contract technology automates the entire transaction and state process, making the steps to be apparent to all users. In the context of the vehicle selling use case, to sell a vehicle, a smart contract can be designed to store all the vehicle-related information in a blockchain database. Similarly, smart contracts are designed with various tasks to facilitate the users of that particular network.

2.1 Smart Contract Attacks

The unaddressed vulnerability and lack of identity anonymity make the well-designed smart contracts, prone to various cyber-attacks on the blockchain platform. Though smart contracts are deterministic, facing issues of nearly 99.9% and approximately 63% of them face critical exposure [5]. Table 1 shows the recently published papers and their contributions to various scenarios.

Table 1 Analysis of attack types

Reference	Contribution	Scenario	Dataset
[6]	SC attacks analyzed and identify vulnerable patterns	Unexpected revert scenario	500 real-time SC from Etherscan
[7]	The static analysis applied to SC distribution and ten common vulnerabilities with remediation suggested	Static analysis	1 million SC
[8]	Preprocess SC by using defender then fuzzy engine generates input and collects execution log of SC	Dynamic analysis	204 SC from ether-scan
[9]	Detecting 8 types of vulnerabilities in terms of higher precision and recall	Static and dynamic analysis	1838 SC
[10]	SC code converted to an XML representation and verify patterns against XPath	Static analysis	Etherscan
[11]	To detect the unwanted behavior of smart contracts, symbolic-based approach was used	Static analysis	165,621 SC
[12]	Smart contract vulnerabilities are identified by using autoencoders and LSTM methods	Dynamic analysis	Etherscan

3 Testing Tools/Techniques

Smart contracts are designed and executed as per the demand of the use cases. Unnoticed vulnerabilities lead to data loss. Before deploying the smart contract, it should be carefully tested with all the possible test cases to catch the vulnerability. Through various analysis tools, smart contract vulnerabilities can be detected. Changes to the smart contract wouldn't be possible, once it is deployed. This section presents some of the benchmarking of the smart contract vulnerability-detecting tools intend for Ethereum. Table 2 displays the testing tools along with the type of attack they can detect. In 2016, the Oyente is an open source tool developed using python at the national university of Singapore. This is the first security analysis tool for analyzing the security breaches in Ethereum smart contracts. Oyente uses the EVM byte code to detect the vulnerabilities which can code written using Solidity 5.0. Oyente acts as a booster for the de velopment of other tools like Maian, Osiris, and Honey-Badger. The tool was tested with 19,366 numbers of Smart Contracts extracted from initial 1,460,000 blocks in the Ethereum network. Oyente is a static analyzer that could detect the possibilities of the latest bugs like DAO, re-entrancy, bad randomness, Call stack vulnerability, unbounded computational power, and GovernMental attack. However, Oyente is not meant for identifying arithmetic vulnerabilities [13].

Slither is an open-source Solidity static analysis tool developed using Python3 which could detect nearly 45 numbers of vulnerabilities. Slither is not only for

Table 2 Smart contract testing tools

Testing tool	Technique	Open source/ language	Attack
Oyente	Static	Y/Python	DAO, re-entrancy, bad randomness, call stack vulnerability, unbounded computational power, GovernMental attack
SmartCheck	Static	Y/Java	DAO, GovernMental attack, Block stuffing attack, Re-entrancy attack, underflow and overflow, unchecked calls, tx.orgin, unchecked math, deprecated, send
Security	Static	Y/Java	Re-entrancy, DAO, GovernMental attack, unchecked calls
Mythril	Static	Y/Python	DAO, Improper Access Control, GovernMental attack, unchecked call, tx.orgin, deprecated, unchecked math, send
vandal	Static and dynamic	Y/python	Re-entrancy, exception disorder, tx.orgin, send, self-destruct
Porosity	Static	Y/c++	Re-entrancy, bad randomness, callstack vulnerability
zeus	Static	N/c++	Re-entrancy
MontiCore	Static and dynamic	Y/python	Re-entrancy, unchecked call, tx.orgin, unchecked math

detecting security breaches but also does the code optimizations that the compiler ignores to address. Slither takes as input the Solidity contract's source code, uses the Solidity compiler to make an abstract syntax tree, and then pulls out the inheritance graph and list of expressions. The Ethereum-based smart contract analysis tool is capable of detecting more than 37 numbers of vulnerabilities. The contract behavior is analyzed with the help of a contract pattern to prevent false negatives. These patterns are evaluated against the semantic facts received from the contact using Souf x001D e Datalog solver which employs static analysis techniques such as data- and control-x001D ow analysis. Mythril was introduced in 2017 by ConsenSys, and it is an open-source analysis tool developed in Python. This tool also accepts EVM byte code as input to perform the analysis. The candidate uses three approaches to analyze smart contracts such as symbolic execution, taint analysis, and SMT solving. Mythril is not only used to analyze the security breaches raised through a smart contract on the Ethereum blockchain but also assists with other blockchains platform-based smart contracts. SmartCheck is an open-source static analyzer tool developed in Java and released in 2017. Smart contracts are analyzed by employing lexical and syntactical methods. A parser and bespoke Solidity language are used to generate an XML parse tree as an intermediate form. Processing of intermediate representation is done using XPath queries to find the vulnerabilities. Through this tool, 20 different types of vulnerabilities can be detected.

4 Use Cases and Future Directions

Smart contract acts as a backbone for every blockchain-enabled use case shown in Fig. 1. One of the reasons for the broader usage of the blockchain platform is its well-drafted deterministic smart contracts. This section enlightens on the importance of smart contracts for implementing various operations about different use cases.

4.1 Smart Contract Based Use Case

Qualities of human lives are greatly improvised as a result of the inclusion of technological advancements. The current decade has greatly benefited through many empowered technologies, among them one such technological advancement is IoT. In the smart healthcare sector, IoT devices are used to gather various health parameters. Gathered data are analyzed at low-end devices and processed internally to get more insight. Processed data saved in the blockchain database to ensure the security aspects. In the traditional database, there is a lack of efficiency in accessing the data without compromising the security aspects. In [14] authors discussed the smart contract for a healthcare management system to share patients' medical records. This mainly focused on the medical prescription handling process to eliminate long waiting times and remove fraud to reduce the error rate by a doctor. Smart contracts

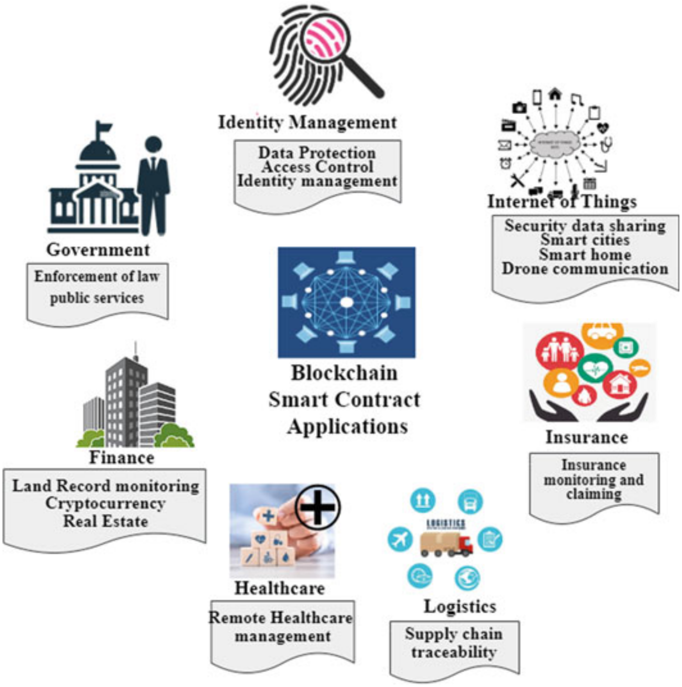


Fig. 1 Blockchain smart contract-based applications

make the system more reliable and automated. Finally, real healthcare datasets are deployed and compare the estimated cost based on various factors.

Vehicular communication: The authors [15] proposed a framework that uses the latest technologies like machine learning and blockchain to find incidents. Machine learning was used to train roadside units (RSU) how to give driving-related warnings based on their location and find problems inside vehicles. With ciphertext policy attribute-based access control (CP-ABE) over a blockchain, access to the incident data was controlled.

Insurance services: Traditional insurance systems consume more time to process a claim, and ambiguity arises among the stakeholders during the process. Using blockchain-based smart contract systems streamlines the process made transparent, secure, and without third-party intervention. Various insurance companies and users register under the blockchain network, and they have a specific address in the blockchain system. The authors [16] present smart contracts based on automatic payments and loss settlement frameworks for insurance services. They consider standard ERC20 token-based smart contracts and Web-based interfaces used to automate crypto-token sales. Smart contract was developed based on three phases, namely analysis, development, and implementation. Finally, the system can handle automatic payments on a claim for a loss that has been approved.

Internet of Things: The Internet of things is a grooming research area. Most IoT devices are resource-constrained and have less power consumption. IoT devices are connected with different applications based on different technologies. For device and data security blockchain integration has been made in smart applications like smart cities, smart homes, and smart transportation. Smart contract-based blockchain integration to become more autonomous for diverse applications and to become more transparent. In [17] IoT, various access control mechanisms have existed like role-based, attribute-based, and capability-based but which are not truly distributed. The authors proposed an access control mechanism by using proof of concept. This allows the decentralized rule evaluation through the device manager. They can add a device, add an access rule, and evaluate that rule based on constraints. However, the system is more scalable compared to existing techniques.

Finance applications: The traditional real estate system faces a lot of issues like database updating, trust among the parties, and disintermediation of stakeholders. A blockchain-based smart contract can reduce the problems connected with real estate. Distributed ledger technology allows all stakeholders to sell and buy their validated and verified documents digitally. These documents are stored inside a ledger that everyone can access. Authors [18] presented a smart contract-based real estate use case. Buyer and seller agreement through a smart contract without involving a third party. Initially, the contract owner creates a function to collect the rent where an agreement was signed by both parties. The payment function is activated, once the payment is over, and then, automatically it will be terminated.

4.2 Challenges and Future Scope

This section presents the possible challenges while designing smart contracts. One of the biggest challenges is the detection of invalid transactions. The problem arises when the numbers of dishonest nodes are higher, and automatically the invalid transaction gets added to the block. Contracts with defects cannot be altered once it is deployed, and hence, more care needs to be given while designing the smart contract. In another aspect, smart contracts need to emphasize the reduction of execution time. Smart contract deployed in Ethereum consumes more time to process and greatly affects network resource. Scalability is the biggest issue associated with blockchain, the inclusion of more users, increases the transaction in turn increases gas consumption. Smart contracts on the Ethereum blockchain are expected to include more features and functionality, and such aspects may increase the possible vulnerabilities as well. To combat the new vulnerabilities, the existing tools need optimization to tackle the new threats. A greater amount of effort might be put toward locating previously unknown vulnerabilities and formulating strategies to mitigate their impact. For example, if you want to be safe if an assault is launched, you can incorporate a self-protection mechanism that is revocable into the smart contract. This is extremely important because, once the smart contract has been deployed on the blockchain, it is unchangeable. Concerns regarding scalability and security are very apparent when the numbers of Ethereum projects continue to expand over time. As a consequence of this, these challenges need to be taken into consideration in the relatively near future concerning the smart contract.

5 Conclusion

Blockchain is the most promising technology due to its amazing features like transparency, immutability, and decentralization. A smart contract is designed to verify and execute tamper-resistant agreements among the respective untrustworthy parties and ensures greater automation. This brief survey enlightens the readers about the various vulnerabilities that need to be focused on while developing smart contracts. And the possible security breaches through unattended vulnerabilities are also furnished. Summarization of various techniques presented various venues in mitigating the security breaches and some of the security breach testing tools along with their properties. Though many countermeasures and testing tools are facilitating in detection and mitigation of security breaches through smart contracts, the candidate problem still needs efficient mechanisms.

References

1. Ramaiah M et al (2023) A review of security vulnerabilities in industry 4.0 application and the possible solutions using blockchain. Cyber Secur Appl Ind 4:63–95
2. Humayun M et al (2021) Internet of things and ransomware: evolution, mitigation and prevention. Egypt Informatics J 22(1):105–117. https://doi.org/10.1016/j.eij.2020.05.003
3. Hutchison D et al (2012) Open problems in network security. IFIPWG 11.4 Int. iN-etSec 2011, vol 7039, p 168. https://doi.org/10.1007/978-3-319-39028-4
4. Wang S et al (2019) Blockchain-enabled smart contracts: architecture, applications, and future trends. IEEE Trans Syst Man Cybern Syst 49(11):2266–2277. https://doi.org/10.1109/TSMC.2019.2895123
5. Tikhomirov S et al (2018) SmartCheck: static analysis of ethereum smart contracts. In: Proceedings international conference on software engineering, October 2017, pp 9–16. https://doi.org/10.1145/3194113.3194115
6. Samreen NF, Alalfi MH (2021) SmartScan: an approach to detect denial of service vulnerability in ethereum smart contracts. In: Proceedings 2021 IEEE/ACM 4th international workshop on emerging trends in software engineering for blockchain, WETSEB 2021, pp 17–26. https://doi.org/10.1109/WETSEB52558.2021.00010
7. Usman TA et al (2021) An analysis of ethereum smart contract vulnerabilities. In: 14th International conference information security and cryptology, ISCTURKEY 2021—Proceedings, no. December, pp 99–104. https://doi.org/10.1109/ISCTURKEY53027.2021.9654305
8. Pan Z et al (2022) ReDefender: a tool for detecting reentrancy vulnerabilities in smart contracts effectively, pp 915–925. https://doi.org/10.1109/qrs54544.2021.00101
9. Akca S et al (2019) SolAnalyser: a framework for analysing and testing smart contracts. In: Proceedings—Asia-Pacific software engineering conference APSEC, vol 2019-December, pp 482–489. https://doi.org/10.1109/APSEC48747.2019.00071

10. Noor Aidee NA et al (2021) Vulnerability assessment on ethereum based smart contract applications. In: 2021 IEEE international conference on automatic control and intelligent systems I2CACIS 2021—Proceedings, June, pp 13–18. https://doi.org/10.1109/I2CACIS52118.2021.9495892
11. Chen J et al (2022) DefectChecker: automated smart contract defect detection by analyzing EVM bytecode. IEEE Trans Softw Eng 48(7):2189–2207. https://doi.org/10.1109/TSE.2021.3054928
12. Narayana KL, Sathiyamurthy K (2022) Automation and smart materials in detecting smart contracts vulnerabilities in Blockchain using deep learning. Mater Today Proc. https://doi.org/10.1016/j.matpr.2021.04.125
13. Kushwaha SS et al (2022) Systematic review of security vulnerabilities in ethereum blockchain smart contract. IEEE Access 10:6605–6621. https://doi.org/10.1109/ACCESS.2021.3140091
14. Khatoon A (2020) A blockchain-based smart contract system for healthcare management. Electronics 9(1). https://doi.org/10.3390/electronics9010094
15. Philip AO, Saravanaguru RAK (2020) Secure incident evidence management framework (SIEMF) for Internet of vehicles using deep learning and blockchain. Open Comput Sci 10(1):408–421. https://doi.org/10.1515/comp-2019-0022
16. Aleksieva V et al (2019) Application of smart contracts based on ethereum blockchain for the purpose of insurance services. In: Proceedings international conference biomed innovations and applications. BIA 2019, pp 18–21. https://doi.org/10.1109/BIA48344.2019.8967468
17. Putra DR et al (2019) Blockchain and smart-contract for scalable access control in Internet of Things. In: Proceeding—2019 International conference ICT Smart Soc Innov Transform Towards Smart Reg ICISS 2019. https://doi.org/10.1109/ICISS48059.2019.8969807
18. Karamitsos I et al (2018) Design of the blockchain smart contract: a use case for real estate. J Inf Secur 09(03):177–190. https://doi.org/10.4236/jis.2018.93013

Automating Fish Detection and Species Classification in Underwaters Using Deep Learning Model

Mamillapalli Chilaka Rao, Praveen Kumar Karri, A. Nageswara Rao, and P. Suneetha

Abstract Ecology is the study of living organisms and checks how they interact with real-world environment. Nowadays, we can see aquatic ecology is increasing its demand in order to maintain very critical information about underwater organisms for efficient conservation and management. As we all know that there is a lot of accessibility, equipments are developed in order to record the underwater footage very accurately compared with manual data collection. Some of the equipments are action cameras capturing 2D and 3D images and unmanned underwater devices, and so on used for finding the details of underwater creatures. Since the human curiosity is increasing more and more, there were a lot of implementations evolved from land to underwater and the space. This is becoming a complicated task for gathering information about dangerous creatures which are residing in underwaters. In current days, robots are information by exploring automating techniques to detect the organism and find the type of species. The use of CNN is to automate fish detection and species identification from underwater videos or images. Normally, we try to apply several ML algorithms to test the species detection, but if the data size is increased, the ML gives very less accuracy. In ML classification algorithms, there is a chance of occurring over fitting problem when the data size is too long; hence based on all these limitations, we try to develop a model using CNN to test the efficiency of our proposed model for accurate fish developed for gathering the ocean detection and species classification from underwater organisms. Our simulations

M. C. Rao (✉)
Department of Information Technology, SRKR Engineering College (A), Bhimavaram, Andhra Pradesh, India
e-mail: chilakarao@gmail.com

P. K. Karri · P. Suneetha
Department of CSE, Sri Vasavi Engineering College (A), Tadepalligudem, Andhra Pradesh, India
e-mail: praveenkumar.cse@srivasaviengg.ac.in

A. Nageswara Rao
Department of CSE, Shri Vishnu Engineering College for Women (A), Bhimavaram, Andhra Pradesh, India
e-mail: anageswararaocse@svecw.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_39

results clearly state that proposed CNN model has achieved 99.39% of training accuracy for performing the given task.

Keywords Deep learning · Convolution neural network (CNN) · Underwater · Kaggle · Unmanned

1 Introduction

Ecology is the study of living organisms and checks how they interact with real-world environment. Marine ecosystem is termed as one form of study of aquatic organisms under deep waters which contains high levels of salt content. Marine ecosystem is categorized into open level ocean, deep sea level and coastal marine level with several characteristics. One among the main level is coastal marine which will provide habitats for spawning, nursing, and feeding for a diverse fish community. As this coastal marine system is very complex and always changes dynamically with environmental changes, it is very difficult to monitor and find out the ecological changes. All the manual approaches failed to achieve proper outcome from this vast coastal environment; hence, this motivated the users to deploy high-resolution underwater video cameras for taking video recordings and capture images from deep underwaters. All the data which is captured will be stored in separate location to cross-check each and every observation and find out the exact species, and fish detection is done from that video or images. In general, the world is divided into sixty-six large coastal ecosystems which contain a vast number of productions compared with oceans. Knausgård et al. [1] adopted the deep learning approach for temperate fish detection and classification. In that article, the authors try to find out and classify each and every fish based on deep learning by taking the help of squeeze-and-excitation (SE) architecture without pre-filtering. Al Muksit et al. [2] adopted the importance of YOLO-Fish detection from realistic underwater environment. In this article, the author concentrated more on the importance of marine research, which is very vast and a lot of species are present still un-named, and he also discussed the importance of several species which are used in medicinal purpose.

Mohamed et al. [3] adopted the method of MSR-YOLO to enhance the fish detection and tracking in the fish farms, in that article the author maximum concentrated on the importance of YOLO in order to classify the fishes in fish farms. Hence, there is very abundant information gathered from all these sources to identify the marine ecosystem and classify how many species are present in this marine ecosystem. All the manual approaches are efficient in identifying some sort of information which are already trained with some clear observations, but manual approaches failed to identify the blur or weak appearance images. Hence, the manual approach is not providing accurate results by classifying with primate machine learning classification techniques. Pagire et al. [4] discussed the importance of deep learning models in order to classify the undersea waters species, and hence, this motivated me to develop a novel deep learning model [4] to detect the coastal ecosystem organisms such as fishes and also try to classify the species of that fish and try to preserve all

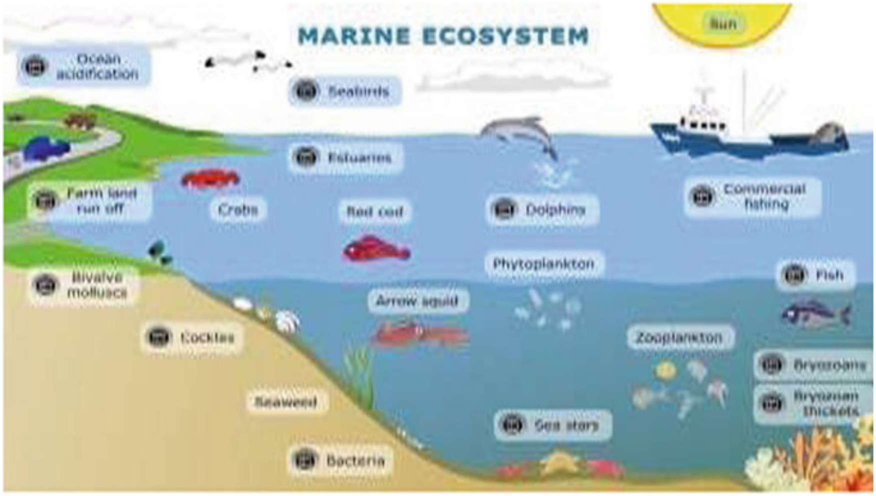


Fig. 1 Flow of marine ecosystem

the useful information under storage media. From Fig. 1, we can clearly identify the life cycle of marine ecosystem describing several living organisms including fishes, crabs, sea stars, bacteria and so on. It is a very complex task for an individual who try to identify each and every organism and classify them with their species names.

2 Literature Survey

Literature survey is the most vital step in the software development process. Before developing the new application or model, it is necessary to work out the time factor, economy and company strength. Once all these factors are confirmed and got approval, then we can start building the application.

Preliminaries

Here in this section, we try to discuss some important attributes which are present in marine ecosystems and then try to explain the importance of marine ecosystems.

Ocean Acidification

This is the primary factor which we need to discuss in marine ecosystems, when the ocean absorbs CO_2, then there will be some changes in ocean. The water chemistry will be changed from normal level to some chemical level, and hence, it causes the pH level of sea to go down. If the pH level goes down, then the water will become more and more acidic which is not good for some living organisms which reside in deep waters. If this type of situation is occurred in marine ecosystems, we need to take preventive steps in order to release less CO_2 [5] and try to reduce the acidic levels of water to give a substantial growth for the underwater organisms.

Sea Birds

These are considered as one of the main sources in food web. If there is no cockles [6] present in life cycle of food web, it will show lot of drastic changes on the consumers who try to consume the ocean food [7]. Because these are the materials used to create the life cycles of the tiny organisms that make up marine ecosystems. We must therefore make sure that the marine habitats are not harmed.

Farm Land Run-Off

This is one of the main sources for growing some plants under downstreams. In order to promote farming, we try to use several fertilizers for cultivation and some sort of over filled fertilizers will reach to underwater carry away through estuaries. If these are done properly, this will help some seaweed to grow very strongly and efficiently under deep waters and this will in turn act as source of food for some fishes and water organisms. And we need to see that some unwanted fertilizers [8] should be avoided for cultivation because if those unwanted fertilizers are entered into underground water levels, then this may harm some living organism and its family.

Estuaries

Estuaries [9] are considered as one of the main sources for all the marine organisms especially for fishes to come at one place in their young stages and perform nursery habits. This will also act like a filter to purify the unwanted materials not to enter into the ocean level.

Fish

This is one of the main organisms which stay in underwaters for feeding a lot of people. These are classified into several categories such as general feeders or special feeders based on the age. If these are not identified properly and categorized, then we cannot able to serve properly for feeding the people, and hence, this should be identified accurately and try to observe whether which species this is categorized [10, 11].

3 Species Classification

In general, the ecosystem contains several types of species, and coming to marine ecosystems, there are several fish species available, and it is a very difficult task for learner to identify the species accurately. For species identification, one should able to learn the life cycle of fish and its entire habitat for classification of fish species [12]. If one can do these things, then only we can able to identify the several kinds of species which are available in the world. The ocean contains several living species which is considered as home of countless organisms living in that vast ecosystems.

Hence, a lot of research projects are going on with several names to detect those organisms and classify them based on species life cycle. For example if we take

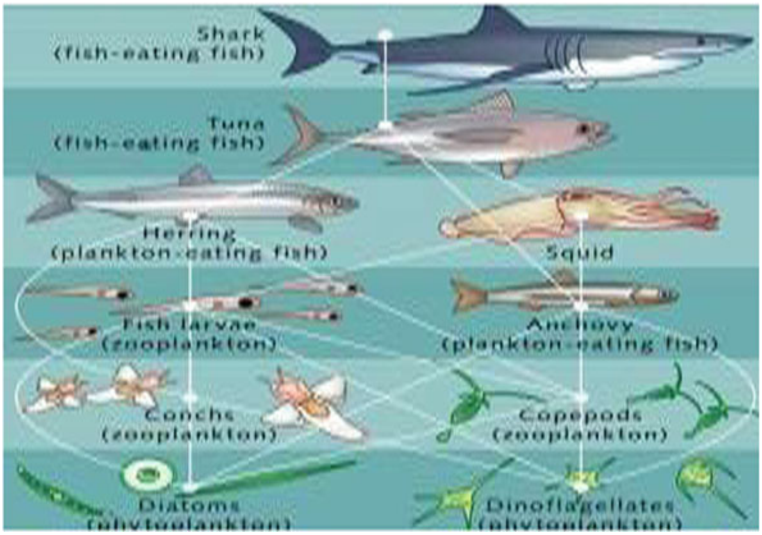


Fig. 2 Marine ecosystem species classification

example of coastal ecosystems, this will contain some normal species to very rich species such as mussels and words which live in Wadden Sea which is food source for several birds. From Fig. 2, we can clearly identify that there are several fishes in this marine ecosystems which will differ one with another. Initially, we try to consider shark which is one of the large fish which can eat other fishes and survive. Next to that we have tuna fish which is also fish-eating fish and so on. This species is classified up to diatoms. So in this proposed application, we try to classify to detect the fishes and identify which species it belongs to and preserve all the information for future predictions.

4 Proposed Dataset

In this proposed work, we try to use **deepfish** dataset which is released in GitHub Web site for research purposes [13]. This dataset contains more than 40 thousand of images collected from deep waters in the marine environments of Australia. In this dataset, more than twenty different habitats are present and all these data we try to train for the CNN model in order to detect the fish and identify which species it belong to.

Dataset is available at https://github.com/DZPeru/fishv3.

From the above source, we can collect the fish dataset, this is almost 700 KB size, and we try to load this dataset into the application for testing and training the proposed CNN model very accurately.

5 Proposed Methodology

In this section, we try to discuss the proposed algorithms which are used for showing the performance of our current objective. In order to prove the performance of our current application, we try to use CNN model which can automatically detect the fishes and identify the type of species based on the sample dataset which is used for training the application. Here we try to divide the model into two parts: one is for fish detection and other is fish species classification [14].

5.1 Fish Detection Phase

In this phase, we try to gather sample input as image file which is collected from marine ecosystem and then loaded as input for training the model. Now this image is passed into hidden layers in CNN model which will be preprocessed and then finally will generate the output whether any fish is detected or not. Here we try to apply max pooling layers, conv layers and dropout layers for performing the desired output. Here we use binary classifiers to divide the image into number of frames. The detection phase mainly uses top-down approach for image search, and then we apply sliding window technique to detect the image in several frames whether any frame is containing a living organism. If there is any frame which is containing a living organism, then it will be identified as fish is present or else it will label as no fish (Figs. 3 and 4).

Fig. 3 Fish detection phase

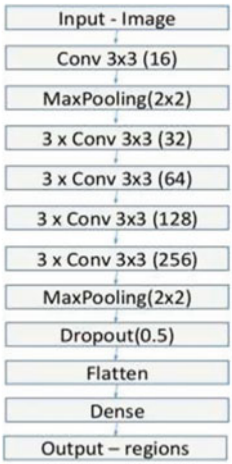
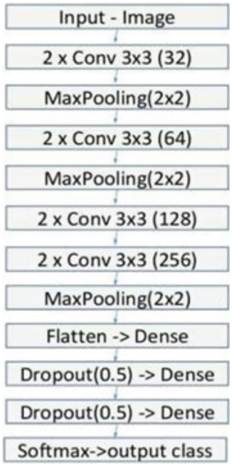


Fig. 4 Classification phase



5.2 *Classification Phase*

Once the image is detected with whether fish is present or not, then we need to search for the type of species which is matched with that species. In the first phase, we try to divide the input image into several frames and each and every frame may be or may not be of same sizes. Hence in this phase, first we try to reconstruct each and every frame into unique size and then pass those frames to classifier network. Here we try to calculate the probability distribution technique to find out whether there is any fish present or not.

Probability equation

$$P(x) = \begin{array}{ll} \max(\mathrm{pi}(x)), & \text{if } x != ' NoF' \\ \mathrm{avg}(\mathrm{pi}(x)), & \text{Otherwise} \end{array}$$

Here we try to use pre-trained CNN model such as VGG-16 for classification of species. The VGG-16 [15] contains several convolutional layers followed by a fully connected (FC) layer. This is mainly formed for classification with the input size of (224, 224, 3). As this is a pre-trained model, we no need to construct the logic from patch onwards; hence, we can use this model according to our application. This model is applied with our input dataset and try to train the model with all the images which are present inside the input dataset. Once the frames which are collected from previous phase are passed to this classification phase as input, this will be undergoing the classification process. In this phase, if the fish is detected accurately in detection phase image, the classifier what we are using in this current model is classified into eight different species, and hence, we can get the result according to the type of

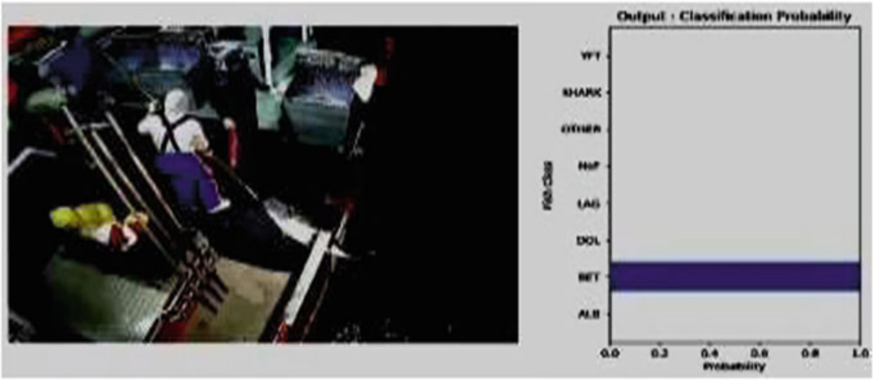


Fig. 5 Classification probability

classifier which is matched with that detected fish. Initially, this fish will be cross-checked with all the eight different levels and those which are having the highest probability will be matched, and this will be mapped accordingly (Fig. 5).

For example in the below frame, we try to give input as one frame which contains the fish, and if that fish is mapped with category BET, then the graph will be corresponding to the BET classifier pointing with high probability.

6 Experimental Results

For showing the performance of our proposed application, we try to deploy the current application using Python as programming language. First we will import all the necessary libraries and then load the input dataset to detect the fish and also find out the type of species it belongs to. Now let us discuss this in detail.

In Fig. 6, we can clearly see there are several libraries and packages used to prove the current objective. Hence, we try to load all those necessary libraries and import them into our application. In Fig. 7, we can clearly see we try to load the fish dataset collected from Kaggle Web site, and this is loaded into the memory for training the application.

Data Cleaning

In this stage, we try to clean the data, all the negative data is removed from the original data source, and only the positive data which is matched with marine ecosystem is kept for training the application. If there are any impurities present in the input data, this will generate **anonymous** result. Hence, we need to clean the data accurately before training the model.

```
[ ] import tensorflow as tf
    from tensorflow.keras.optimizers import RMSprop
    from tensorflow.keras.preprocessing.image import ImageDataGenerator
    from tensorflow.keras import layers
    from tensorflow.keras import Model
    import matplotlib.pyplot as plt
    import os
    import zipfile
    import numpy as np
    from google.colab import files
    from keras.preprocessing import image
```

Fig. 6 Import libraries

```
import random
from shutil import copyfile

# The following python code will use the OS library to use Operating System libraries
# and giving  access to the file system, and the zipfile library allowing to unzip the data.

local_zip = '/content/data/QUT_fish_data.zip'

zip_ref = zipfile.ZipFile(local_zip, 'r')
zip_ref.extractall('/content/data/')
zip_ref.close()

# Use os.mkdir to create your directories

try:
    os.mkdir(' /content/data/images/training (ctrl + click)
    os.mkdir('/content/data/images/training')
    os.mkdir('/content/data/images/testing')
    os.mkdir('/content/data/images/training/pos')
    os.mkdir('/content/data/images/testing/pos')
except OSError:
    pass
```

Fig. 7 Load input dataset

From Fig. 8, we can clearly identify fish is detected and normal objects are identified as no fish, and we can test on any image as sample input and our application will automatically identify whether any fish is present or not.

From Fig. 9, we can clearly identify the species of that fish and its highest probability.

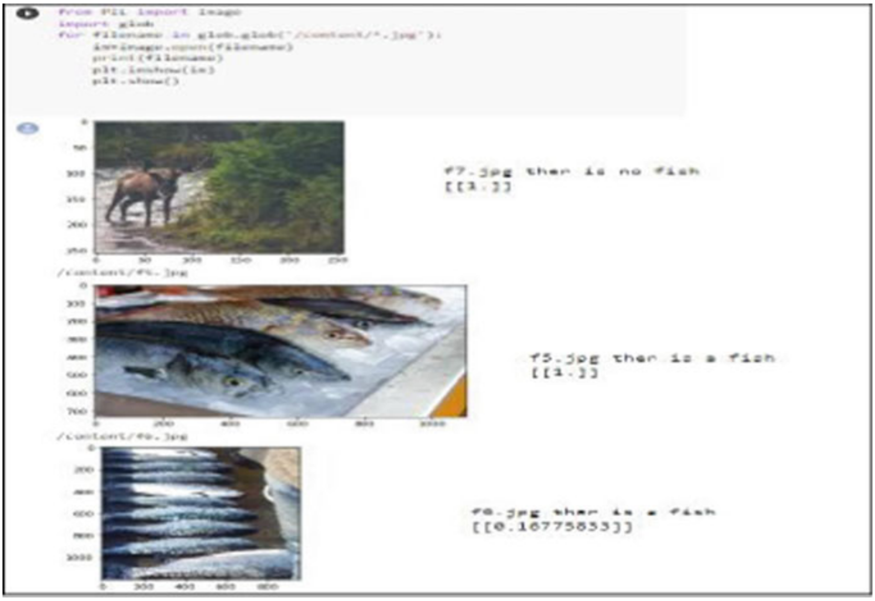


Fig. 8 Fish identification

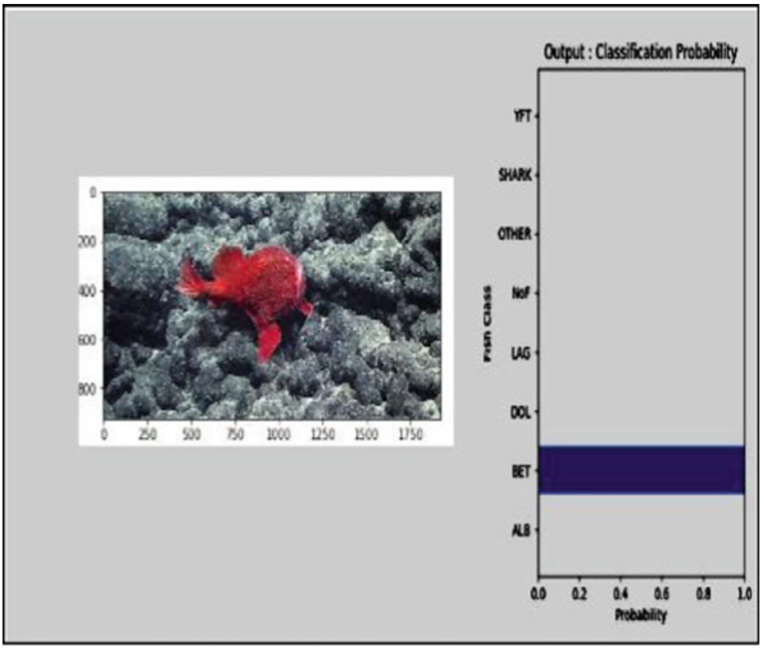


Fig. 9 Species classification graph

7 Conclusion

In this proposed work, we have tested a large number and variety of different fish species collected from coastal area of Australia and then finally applied CNN model for detection and classification. By performing several theoretical and experimental analyses, we finally got the best performance with nearly 99.39 % accuracy by using CNN model. In general, I strongly believe that this research is very helpful for coastal ecosystem for finding several new organs and name them with corresponding species so that we can show their existence for this upcoming universe. In future, I want to continue the same work for my research and concentrate on data anonymous functionality so that if there is any anonymous data present in the input dataset, those should be automatically identified and removed in data cleaning stages.

References

1. Knausgård KM, Wiklund A, Sørdalen TK et al (2022) Temperate fish detection and classification: a deep learning based approach. Appl Intell 52:6988–7001. https://doi.org/10.1007/s10489-020-02154-9
2. Al Muksit A, Hasan F, Emon MFHB, Haque MR, Anwary AR, Shatabda S (2022) YOLO-Fish: a robust fish detection model to detect fish in realistic underwater environment. Ecol Inform 72:101847. ISSN 1574-9541
3. Mohamed HED, Fadl A, Anas O, Wageeh Y, ElMasry N, Nabil A, Atia A (2020) MSR-YOLO: method to enhance fish detection and tracking in fish farms. Procedia Comp Sci 170:539–546. ISSN 1877-509
4. Pagire V, Phadke A (2022) Underwater fish detection and classification using deep learning. In: 2022 International conference on intelligent controller and computing for smart power (ICICCSP), Hyderabad, India, 2022, pp 1–4. https://doi.org/10.1109/ICICCSP53532.2022.9862410
5. https://www.sciencedaily.com/releases/2021/02/210217132320.htm
6. https://www.ocean.org/blog/trophic-levels-part-1/
7. Therry L (2022) Effects of fisheries on marine biodiversity. Available from http://www.coastalwiki.org/wiki/Effects_of_fisheries_on_marine_biodiversity
8. Rahman KMA, Zhang D (2018) Effects of fertilizer broadcasting on the excessive use of inorganic fertilizers and environmental sustainability. Sustainability 10:759. https://doi.org/10.3390/su10030759
9. Tendolkar A et al (2021) Estuarine-Star (smart estuarine aquaculture management with sensor network). In: 2021 IEEE international conference on systems, man, and cybernetics (SMC), Melbourne, Australia, 2021, pp 1593–1598. https://doi.org/10.1109/SMC52423.2021.9659122
10. Li X, Shang M, Qin H, Chen L (2015) Fast accurate fish detection and recognition of underwater images with Fast R-CNN. In: OCEANS 2015—MTS/IEEE Washington, Washington, DC, 2015, pp 1–5. https://doi.org/10.23919/OCEANS.2015.7404464
11. Szegedy C, Liu W et al (2015) Going deeper with convolutions. In: CVPR. Google Research
12. Rathi D, Jain S, Indu S (2017) Underwater fish species classification using convolutional neural network and deep learning. In: 2017 Ninth international conference on advances in pattern recognition (ICAPR), Bangalore, India, 2017, pp 1–6. https://doi.org/10.1109/ICAPR.2017.8593044
13. https://github.com/DZPeru/fishv3

14. Dewan J, Gele A, Fulari O, Kabade B, Joshi A (2022) Fish detection and classification. In: 2022 6th International conference on computing, communication, control and automation (ICCUBEA, Pune, India, 2022, pp 1–5https://doi.org/10.1109/ICCUBEA54992.2022.10010836
15. Kholoud KKB, Elbatsh K, Sokar I, Ragab S (2021) WikiFish: mobile app for fish species recognition using deep convolutional neural networks. In: 2021 The 4th international conference on computational intelligence and intelligent systems (CIIS 2021), November 20–22, 2021, Tokyo, Japan. ACM, New York, NY, USA, 10p. https://doi.org/10.1145/3507623.3507626

Explorative Analysis of Factors Influencing Alcohol Risk Level Based on Epidemiological Data and Prediction Using Machine Learning Techniques

P. G. Sunitha Hiremath, Shobhit Nigam, K. Kartik, Sumit Sharan, G. Suraj, and Vivek V. Pais

Abstract Many factors influence alcohol risk drinking in individuals, so to explore the same and to predict it, we considered Epidemiological data of 2972 male patients' from a de-addiction centre, in South India. In this data, 79% patients were in very high risk, 17.56% in high risk, 0.72% in medium risk and 2.70% in low risk level. Among 75 features, 35 features were selected as relevant using Extra Tree Classifier feature selection method. Explorative data analysis was carried out to understand the relationship between various features and risk level. The class imbalance in the data was overcome by using Synthetic Minority Oversampling Technique. The Random Forest classifier outperformed when compared to other models in predicting alcohol risk level with an accuracy of 92.84%.This is our preliminary work which has given most promising results and can be used to assist alcohol risk level diagnosis in hospitals or de-addiction centres.

Keywords Risk level · Features · Epidemiology · Adverse child hood experience · Classifier

1 Introduction

Alcohol is one of the easily accessible legal drug used by people worldwide. In developing country like India, the percentage of people consuming alcohol has noticeably increased [1]. The age at which people start drinking alcohol is varying, and it depends on many factors, but the more vulnerable age is the early adulthood [2, 3], for which people give many reasons, like peer pressure or to cope with stress [4].

P. G. Sunitha Hiremath (✉) · S. Nigam · K. Kartik · S. Sharan · G. Suraj
KLE Technological University, Hubballi, India
e-mail: pgshiremath@kletech.ac.in

V. V. Pais
SDM De-Addiction and Research Centre, Ujire, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_40

But, if their alcohol consumption reaches high or very high risk level, then definitely it will have an impact on the physiological and/or psychological health [5–7]. World Health Organization has defined the risk levels as Low, Medium, High and Very High Risk that corresponds to 1–40 g, 41–60 g, 61–100 g and more than 101 g of ethanol consumption per day for males [8]. People who have developed risk drinking problem usually take treatment as an inpatient in hospital or de-addiction centres. We wanted to explore features of such patients from the data and build an alcohol risk level predictive model, which can act as risk diagnostic assistance tool in hospitals and/or de-addiction centres.

In Sect. 2, we discuss about the data collection and its details; pre-processing and data analysis are discussed in Sect. 3; in Sect. 4, model building and evaluation is being discussed followed by conclusion.

2 Methodology

To carry out this study, we collected 2972 male patients' data who took treatment for their risk drinking problem, from 16-04-2019 to 09-03-2020 in Shri Dharmasthala Manjunatheshwara De-addiction and research centre, Ujire, a place in Karnataka state in India. The commonly used methods for assessing the alcohol drinking problem are a questionnaire related to drinking habits, physical examination, laboratory tests and imaging tests, psychological evaluation, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria. In the SKDRDP De-addiction centre, the above listed methods are used to assess alcohol drinking problem except the laboratory and imaging tests. The patients were counselled during their treatment in the centre by trained counsellors, treated by nurses, physician and psychiatrist, who collect the information related to their socio-demographic details, drinking initiation and pattern, withdrawal symptoms, causes and impacts. A written consent of approval is taken from them for using their data for research purpose.

The details of the patients which we considered for alcohol risk prediction are (1) the details about drinking initiation and pattern that include age of first use, years of use, years of excessive use, quantity used in last 30 days and type of alcohol, (2) withdrawal symptoms experienced, (3) reasons to start and continue the consumption of alcohol, (4) chronic health problems, (5) psychiatric complications, (6) adverse childhood experiences, (7) behavioural problem identified in childhood/adolescence, 8) nicotine use: smoking/smokeless, (9) age, (10) education, (11) weight at the time of admission and discharge, (12) stressors, (13) family history of alcoholism, (14) living arrangement, (15) family history of psychiatric illness, (16) marital history and (17) occupational history. The proposed system model for risk level prediction is shown in Fig. 1.

The system model consists of following components:

1. **Data Pre-processing**: Converted all the categorical data into numerical data and filled missing values with the median values.

Fig. 1 System model for risk level prediction

2. **New Features Generation**: Generated three new features alcohol risk level, alcohol withdrawal syndrome (AWS) stages and adverse childhood experiences (ACE) score.
3. **Feature Extraction**: Extracted 35 important features using Extra Tree Classifier.
4. **Data Imbalance**: Data was balanced using Synthetic Minority Oversampling Technique (SMOTE).
5. **Model Building**: Random Forest, Decision Tree and KNN classifiers used for prediction.
6. **Model Evaluation**: Models evaluated using accuracy, precision, recall and F1-score.

The next section discusses about the pre-processing of the data that was carried out to get quality data, exploratory analysis to understand the relationship between the features, model building and its evaluation which states how much accurately the model predicts the risk level for test data.

3 Experiments and Results

3.1 Pre-processing

In this data, 1.74% of patient's records which had multiple missing field values were eliminated. The other numerical field values which were missing were filled with statistical measure mean. The unwanted text which were tampered with the numerical values was removed; for example the value of the attribute "duration of excessive use of alcohol" need to be numerical like 25, but in some fields it was written as 25 years. For few fields more than one values are to be entered and in the dataset, it was recorded in the text format separated by commas. For example, the 'withdrawal symptoms experienced when the patient stopped' field can have more than one value, i.e. a patient may have a set of withdrawal symptoms. Therefore, for every unique withdrawal symptoms a new field or a feature was created which takes either 1 [presence of withdrawal symptom] or 0 [patient do not experience it] using one hot encoding method. Same procedure was followed for other fields like reasons to start, reasons to continue, ACE, problematic childhood/adolescent behaviour, chronic health issues and psychiatric issues, stressors and living arrangements. Thus, the values of the attributes have taken the values either 0 or 1 [9].

The withdrawal symptoms are some physiological and/or psychiatric problems experienced by people when their alcohol intake was suddenly stopped and may experience more than one symptom. These symptoms are categorized in to four stages (attributes) as Alcohol Withdrawal Stage; AWS 1-AWS 4 based on their severity [10]. So these attributes will take the value as 1 if the symptoms are present else 0. A new attribute risk level was generated based on quantity of alcohol consumption. It was calculated as product of average quantity of alcohol consumption, percentage of ethanol content and alcohol conversion factor (1 ml ethanol = 0.79 g) [8]. This attribute has four values based on grams of ethanol consumption for males: Low Risk (1–40), Medium Risk (41–60), High Risk (61–100) and Very High Risk (> 101). Thus, a total of 75 features were considered for our study and given as input to the Extra Tree Classifier. This classifier extracted 35 relevant features and listed them in the decreasing order of their importance as shown in Fig. 2.

The attribute risk level is considered as class label with values Low Risk (1), Medium Risk (2), High Risk (3) and Very High Risk (4). But there was a class

```
['AAO for alcohol in years',
 'Age',
 'Weight while admission (In Kg)',
 'duration of use of alcohol',
 'At what age did you start working?',
 'duration of excessive use of alcohol',
 'AWS_Stages',
 'S_Family or relationship issues',
 'ACE_Poverty or severe debts',
 'S_Financial Stress',
 'S_Work related stress',
 'Nicotine (yes/NO)',
 'Marital Status',
 'R_to feel better/confident/happy',
 'R_to try',
 'R_to avoid problems and sadness',
 'smoking/smokeless',
 'At present do you have any sexual problem ( if yes mention)',
 'Family history of alcoholism / drug abuse, if any (who and which type of drug)',
 'S_Reports Stressed but doesn’t know where or what',
 'How many first degree relatives had Substance addiction',
 'Legal complications yes/no',
 'Record extra marital experiences',
 'any instance of family violence',
 'ACE_Scholastic backwardness',
 'multiple marriages',
 'C_Diabetes',
 'Did you have any period of unemployment',
 'ACE_Early parental loss',
 'Living arrangement_Family',
 'Psy_Confusion',
 'Psy_Hallucinations',
 'Psy_Aggressive Outbursts',
 'ACE_Broken home or single parenting',
 'ACE_Running away from home']
```

Fig. 2 Relevant features

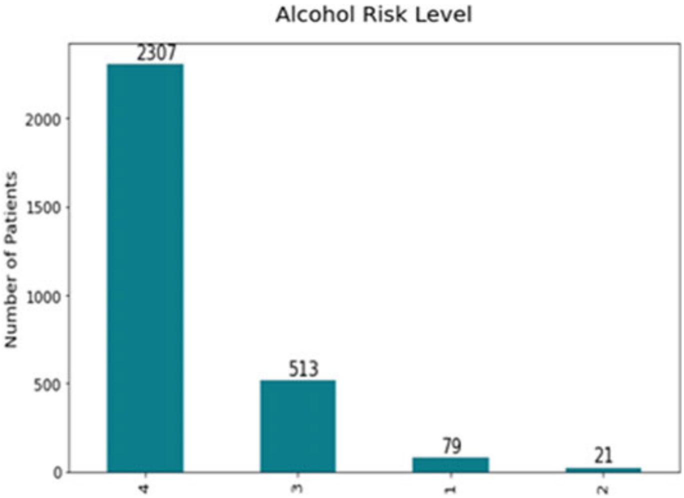


Fig. 3 Number of patients in alcohol risk level technique

imbalance in the data. The Synthetic Minority Oversampling Technique (SMOTE) technique was used for handling class imbalance as shown in Fig. 3. Here, synthetic samples are generated for the minority class to equal the number of records in all the four classes. Thus after applying SMOTE, 2307 patient records are obtained with each class labels, and in total, 9228 patient records were obtained as shown in Fig. 4. The next section discusses about the exploration of relationship between various predictor variables and alcohol risk drinking level.

3.2 Explorative Analysis of the Data

In this section, the relationship between prominent predictor variables and alcohol risk drinking level is being discussed. The patients were categorized in to different age groups as given in Table 1. The other details of the patients correspond to number of patients in each age group, their age, weight, average quantity and duration of alcohol consumption. The values of the attributes are represented as mean ± standard deviation.

As given in Table 1, maximum 74% patients are in the age range of 26–46 years. The average quantity of alcohol consumption is seen maximum in 26–32 age group. In the later age groups especially from 47 to 53 onwards, it is decreased even though duration of consumption is increased. Therefore, to investigate its relationship to risk drinking statistical measure correlation was computed. A weak negative correlation of 0.052 was observed between age and quantity of alcohol consumption of patients, which matches with the most of the works [11, 12]. So age can be considered as one of the prominent predictor variables. The most common reasons for decrease in

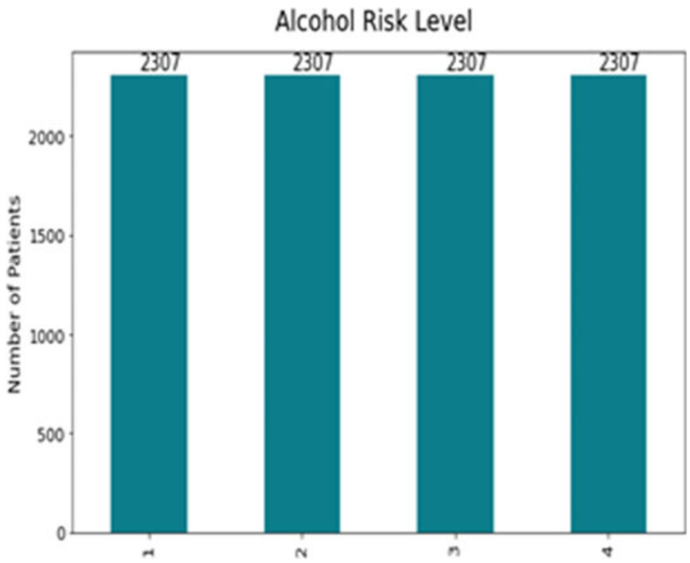


Fig. 4 Number of patients records after applying SMOTE

Table 1 Patients' details

Age group	#patients	Age (in years)	Weight (kg)	Duration of consumption (Years)	Average quantity of Consumption (ml)
18 to 25	190	23.02 ± (1.89)	58.64 ± (10.50)	5.55 ± (3.90)	2434.81 ± (271.87)
26 to 32	714	29.46 ± (1.89)	60.37 ± (12.06)	8.93 ± (6.22)	7329.75 ± (1080.16)
33 to 39	821	35.88 ± (1.94)	61.87 ± (12.18)	12.47 ± (6.00)	495.69 ± (325.91)
40 to 46	612	42.72 ± (2.0)	61.88 ± (11.71)	16.80 ± (7.07)	617.67 ± (364.94)
47 to 53	338	49.55 ± (1.88)	62.40 ± (12.00)	22.79 ± (15.49)	431.06 ± (283.78)
54 to 60	157	56.97 ± (2.15)	60.94 ± (11.45)	26.21 ± (10.96)	422.67 ± (249.62)
Above 60	88	65.25 ± (3.69)	58.50 ± (13.73)	23.73 ± (15.68)	341.64 ± (279.59)

the quantity of consumption of alcohol at the later ages may be people start taking precautions of their health, social occasions and cutting down the cost of drinking [13, 14]; however, this information was not recorded in our data. Many works reported an association between heavy drinking and weight gain in people [15]. This association is more seen in males than females [16]. Males with higher body mass index (BMI) drink more when compared to the lower BMI males [17]. The reason for the weight gain is that the heavy alcohol consumption may contribute to the excess energy intake which is associated with weight gain in some individuals [18]. In this data, a weak positive correlation of 0.051 was observed between risk drinking patients and their weight. A marginal increase in weight was observed in patients of age up to 53 years and in the later age group onwards it's decreased. Thus, a conflicting outcome of decrease in weight and quantity of alcohol consumption is observed in age groups greater than 53 years.

Adverse childhood experiences (ACE) are some unpleasant events experienced by people in their childhood. In this data, there are six categories of ACEs are present; they are poverty/severe debts, early parental loss, extra marital affairs of parents, broken home/single parenting, sexual abuse and violence. These categories of ACEs are associated with risk drinking in adulthood [19]. It was found that 27.97% of total patients experienced poverty/severe debts, 5.1% with early parental loss, 5.92% with single parenting and 5.56% with violence. The effects of other two categories of ACE were not significantly seen in the data. Thus, the most significantly seen category of ACE was poverty/severe debts. There is a possibility that patients might have experienced more than one category of ACE. So ACE score was computed for each patient based on the number categories of ACEs they experienced with. A highest ACE score of 5 was observed, which was seen only in one patient. There was a weak positive correlation of 0.012 between ACE score and risk drinking. That means single ACE may also have an impact than the cumulative effect.

Similarly, the categories of problematic behaviours seen in the patient's childhood/ adolescent ages were investigated. It's observed that 6.9% patients reported running away from home, 9.75% with scholastic backwardness, 6.2% with alcohol use, but very less percentage of patients reported with other factors like stealing, gambling, sexual issues, frequent physical fights and violence, destruction of others property, experimenting with drugs. These ACEs and childhood behavioural problems had an impact on the education level of the patients. Around 40% of the patients had attained maximum of secondary school level and remaining lower than this. The study did not investigate further the relationship of education with occupation and income which is the limitation. But the prominently seen reasons to initiate and continue consumption of alcohol are to try, to feel better/confident/happy and to avoid problems/sadness.

Similarly, the relationship between other factors with risk drinking is found out using correlation coefficient. A positive correlation of work related stress of 0.0012, financial stress of 0.007 and period of unemployment of 0.006 was observed with risk drinking, but a weak negative correlation of 0.0019 was observed between family or relationship issues with risk drinking. These relationships show that work related and /or financial stress might have driven patients towards risk drinking than family issues. This matches with the previous analysis, which showed that maximum quantity of alcohol consumption was seen in the age group of 26–32 years, in which people in India try getting in to good jobs. Further, the analysis explored a positive correlation between family history of alcoholism (0.032) with risk alcohol drinking which matches with the findings in [14], with one additional point which states that the age of onset of initiation is a better predictor of severity of alcoholism, than family history of alcoholism alone. The age of onset of alcohol initiation on risk drinking will be explored in our future work. A positive correlation of 0.033 between family psychiatric illness and risk drinking shows that family problems like psychiatric illness may also lead to risk drinking. The next section discusses the model building and evaluation.

4 Risk Level Prediction Model Building and Evaluation

The data set consists of 2920 patients' records with 2307 in very high risk class, 513 in high risk, 21 in medium risk class and 79 in low risk class. The SMOTE technique was applied to balance this data, and as a result, 9228 patients' records were generated with equal number of records in each class. The data was divided into training and test data in the ratio 80:20. The number of features considered was 35 which were extracted using Extra Tree Classifier. The model was built with different machine learning techniques like Random Forest, Decision Tree and K-Nearest Neighbours to observe which could provide the highest predictive accuracy. The performance of the classifier models are measured using confusion matrix. Table 2 shows the confusion matrix of Random Forest classifier. The number of tuples belonging to each class is almost same ranging from 435 to 495.The macro-average and weighted average for each of the metrics precision, recall and F1-score are same, with an overall accuracy of the model in predicting the Risk Level is of 93%. The performance of all the three algorithms is given in Table 3.

Table 2 Confusion matrix of Random Forest classifier

	Precision	Recall	F1-score	Support
	0.97	0.98	0.97	444
	1.00	0.99	0.99	472
	0.89	0.82	0.86	435
	0.85	0.91	0.88	495
Accuracy			0.93	1846
Macro average	0.93	0.92	0.93	1846
Weighted average	0.93	0.93	0.93	1846

Table 3 Performance comparison of models

Model	Accuracy	Precision	Recall	F1-score
Random Forest	0.928494	0.940525	0.939687	0.940136
Decision Tree	0.864868	0.862597	0.864104	0.863138
K-Nearest Neighbours	0.775673	0.778863	0.775169	0.774894

It can observed from Table 3 that the Random Forest classifier has outperformed followed by Decision Tree and K-Nearest Neighbour. The experiment was performed maintaining the N_estimators to 800 and cross-validation set to 10. The precision is 94%, and recall is 93.9%, which signifies the model is able to classify the risk levels to a satisfactory extent. Precision–recall being high concludes that the model is accurately classifying the class although there is an imbalance in class observed. We are the first one to carry out this kind of risk drinking level prediction; hence, we could not compare the performance of our model with others work.

5 Conclusion

Risk drinking was observed more in the 26–32 years of age, and a decline was seen in the later age groups. Patients with risk drinking had work related, financial and/or unemployment stress which is shown by their positive correlation with risk drinking level. A weak positive correlation between age and weight was observed but after 53 years of age, there was a decrease in the mean weight of the patients and also the consumption of alcohol. Even though with age, weight and also quantity of consumption of alcohol has decreased, the patients were in high or very high risk level. The prominently seen ACE and problematic childhood/adolescent behavioural factors in the data are poverty and scholastic backwardness. These factors affected patient's development during their childhood, their education and drove them towards substance use like alcohol. Family related issues like psychiatric illness and alcoholism shown a positive correlation with risk drinking level. Thus from this study, we could find the prominent influencing factors of risk drinking and built a better risk predictive model. Random Forest performed better with an accuracy of 92.84% when compared to other models. This kind of risk level drinking prediction using de-addiction patient's data is not done before.

The strength of our work is we are working with de-addiction centre's patient data set which consists of all the details required for the analysis. The limitation of this work is that we did not consider many of the other patient's details related sexual history, the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) details, maximum period of abstinence and possible factors for abstinence, previous period of treatment and period of sober, reasons for relapse which might have contributed much to the predictive model. Our future work is to build a Web application for patients' data management and integrate risk predictive model, which can be used as an assistance tool for the personnel at the de-addiction centre. It can also be used to educate patients to assess their risk level.

Acknowledgements We thank Dr. D. Veerendra Heggade, President "Shri Dharmasthala Manjunatheshwara De-addiction and Research Centre", Ujire, Karnataka State, India, for having given permission to carry out research and insightful discussion about alcoholism, its effects and prevention.

References

1. Ambear A, Agrawal A, Rao R, Mishra AK, Khandelwal SK, Chadda RK on behalf of the group of investigators for the National Survey on Extent and Pattern of Substance Use in India (2019) Magnitude of substance use in India. Ministry of Social Justice and Empowerment, Government of India, New Delhi
2. Kramer JR, Chan G, Dick DM, Kuperman S, Bucholz KK, Edenberg HJ, Polgreen LA, Hesselbrock VM, Schuckit MA, Nurnberger JI, Kapp ES, Porjesz B, Bierut LJ (2008) Multiple-domain predictors of problematic alcohol use in young adulthood. J Stud Alcohol Drugs 69(5):649–659. https://doi.org/10.15288/jsad.2008.69.649. PMID: 18781239; PMCID: PMC2575389
3. Petit G, Kornreich C, Verbanck P, Cimochowska A, Campanella S (2013) Why is adolescence a key period of alcohol initiation and who is prone to develop long-term problem use? A review of current available data. Socio Affect Neurosci Psychol 3:1, 21890. https://doi.org/10.3402/snp.v3i0.21890
4. Abbey A, Smith MJ, Scott RO (1993) The relationship between reasons for drinking alcohol and alcohol consumption: an interactional approach. Addict Behav 18(6):659–670. https://doi.org/10.1016/0306-4603(93)90019-6
5. Johannessen EL, Andersson HW, Bjørngaard JH, Pape K (2017) Anxiety and depression symptoms and alcohol use among adolescents—a cross sectional study of Norwegian secondary school students. BMC Public Health 17:494. https://doi.org/10.1186/s12889-017-4389-2
6. Li H, Lv J, Yu C, Guo Y, Bian Z, Fan J, Yang L, Chen Y, Du H, Long H, Zhang Z, Chen J, Chen Z, Huang T, Li L (2020) China Kadoorie Biobank collaborative group. The association between age at initiation of alcohol consumption and type 2 Diabetes Mellitus: a cohort study of 0.5 million persons in China. Am J Epidemiol 189(12):1478–1491. https://doi.org/10.1093/aje/kwaa119. PMID: 32875324; PMCID: PMC7857693
7. Simet SM, Sisson JH (2015) Alcohol's effects on lung health and immunity. Alcohol Res 37(2):199–208. PMID: 26695745; PMCID: PMC4590617
8. World Health Organization (2000) International guide for monitoring alcohol consumption and related harm. World Health Organization. https://apps.who.int/iris/handle/10665/66529
9. Mallya GP, Hiremani A, Gajakosh G, Badiger D, Sunitha Hiremath PG (2018) Framework for automatic diagnosis of psychological disorders. In: 2018 International conference on advances in computing, communication control and networking (ICACCCN), pp 929–934. https://doi.org/10.1109/ICACCCN.2018.8748450
10. Mirijello A, D'Angelo C, Ferrulli A, Vassallo G, Antonelli M, Caputo F, Leggio L, Gasbarrini A, Addolorato G (2015) Identification and management of alcohol withdrawal syndrome. Drugs 75(4):353–365. https://doi.org/10.1007/s40265-015-0358-1
11. Higuchi S, Parrish KM, Dufour MC, Towle LH, Harford TC (1994) Relationship between age and drinking patterns and drinking problems among Japanese, Japanese-Americans, and Caucasians. Alcohol Clin Exp Res 18(2):305–310. https://doi.org/10.1111/j.1530-0277.1994.tb00018.x. PMID: 8048731
12. Chaiyasong S, Huckle T, Mackintosh A-M, Meier P, Parry CDH, Callinan S, Viet Cuong P, Kazantseva E, Gray-Phillip G, Parker K, Casswell S (2018) Drinking patterns vary by gender, age and country-level income: cross-country analysis of the international alcohol control study. Drug Alcohol Rev 37:S53–S62. https://doi.org/10.1111/dar.12820
13. Britton A, Bell S (2015) Reasons why people change their alcohol consumption in later life: findings from the Whitehall II Cohort study. PLoS ONE 10(3):e0119421. https://doi.org/10.1371/journal.pone.0119421
14. Chartier KG, Thomas NS, Kendler KS (2017) Interrelationship between family history of alcoholism and generational status in the prediction of alcohol dependence in US Hispanics. Psychol Med 47(1):137–147
15. Sung K, Kim S, Reaven G (2007) Relationship among alcohol, body weight, and cardiovascular risk factors in 27,030 Korean men. Diabetes Care 30:2690–2694. https://doi.org/10.2337/dc07-0315

16. Arif AA, Rohrer JE (2005) Patterns of alcohol drinking and its association with obesity: data from the third national health and nutrition examination survey, 1988–1994. BMC Public Health 5(5):126. https://doi.org/10.1186/1471-2458-5-126.PMID:16329757;PMCID:PMC1318457
17. Nies MA, Sun L, Kazemi D, Carriker A, Dmochowski J. Relationship of body mass index to alcohol consumption in college freshmen. Scientific World J 2012:849018. https://doi.org/10.1100/2012/849018. Epub 2012 May 2. PMID: 22654639; PMCID: PMC3354654
18. Traversy G, Chaput JP (2015) Alcohol consumption and obesity: an update. Curr Obes Rep 4(1):122–130. https://doi.org/10.1007/s13679-014-0129-4.PMID:25741455;PMCID:PMC4338356
19. Kauhanen L, Leino J, Lakka HM, Lynch JW, Kauhanen J (2011) Adverse childhood experiences and risk of binge drinking and drunkenness in middle-aged Finnish men. Adv Prevent Med 2011:12. Article ID 478741. https://doi.org/10.4061/2011/478741

A Real-Time Hadoop Big Data Maintenance Model Using a Software-Defined and U-Net Deep Learning Mode

Yallapragada Ravi Raju and D. Haritha

Abstract An advanced big data platform can be offered many applications like cloud, server, and Hadoop maintenance. The available big data platform with Hadoop modeling cannot provide dynamic actions in quick time. The latency and storage issues are diminishing the application's robustness. Big data platform and the Internet has also not been adequately used to direct the manufacturing process in cloud storage maintenance. Big data cloud systems will benefit from continual development and evolution guided by big data handled at the gateways in the cloud. This study also presents the latest method (DL-enabled operations Facilities) through a sponsoring software arrangement, this may encourage quick procedure and updating of big data file maintenance techniques with intelligent video surveillance in a closed loop. This work can help cloud-based big data file production systems quickly and operate more efficiently. In this research work U-net based on Hadoop and sparks, big data analysis is performed. The following application is verified on the python 3.7 software tool. The performance measures like accuracy 98.34%, recall 97.23%, sensitivity 98.92%, and throughput 99.23% had been attained which are outperformance the methodology. This U-net-based big data analytics application is competing with present technology.

Keywords Big data · U-net · Deep learning · Hadoop · Sparks

Y. R. Raju (✉) · D. Haritha
Department of Computer Science and Engineering (CSE), KoneruLakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India
e-mail: ravi.y40@gmail.com

D. Haritha
e-mail: haritha_donavalli@kluniversity.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_41

1 Introduction

The cloud applications come to improving and implementing a broad range of smart devices and services to construct so-called smart factories, and manufacturers have been working to shrink the size and cost of microsensors while simultaneously reducing power consumption and improving accuracy. IoT and 5G technologies are often used to link smart items and assets with ID, sensing, and actuation capabilities [1] and easily incorporate them into smart industrialized frameworks like cloud manufacturing (Mfg) systems [2] and Industrial IoT (IoT) technologies. Despite businesses' claims of physical and intangible advantages, they have acquired a great quantity of real-time data from shop floors or marketplaces, social networks, and other sources, including employees. As a result, the industrial industry has to make better use of huge data in order to enjoy the benefits of analytics and artificial intelligence. Service-oriented manufacturing (CMfg) allows for a vast shared resource pool of customizable equipment, networking, and computing (e.g., network and server resources) that can be swiftly delivered or released [2]. This allows for a highly elastic capacity for managing manufactured big data. Moreover, effective techniques to quickly take related actions should be investigated when certain information or insights are mined from big data to boost productivity or decrease losses. Time-sensitive shop floor applications would not be able to meet their needs with a cloud-based data connection, data synchronization, and calculations.

2 Literature Survey

Shi et al. [3] explains about wired communication methods, such as field buses and specialized industrial networks, which are still utilized in the manufacturing sector, while transmitters such as WiFi and ZigBee have lately been embraced. 5G wireless transmission technology integration into cooperative intelligent manufacturing (CIM) processes and systems is the subject of this research project, which will look at it. This study examines 5G technologies in an effort to link manufacturing equipment, smart goods, logistics systems, and information assurance in an effective manner. For heterogeneous wireless convergence, autonomous data collecting, data analysis, and more, we expect CIM systems to include 5G technology in the near future, giving them a promising future [4].

Thus, we offer a demonstration of product design and manufacture an IoT-enabled cloud industrialized atmosphere that is fully connected. Open innovation may be facilitated by the use of social networks in conjunction with the Internet of Things, cloud manufacturing, and different elastic services, such as an on-demand workspace or knowledge sharing or communal issue solving.

It's no secret that CMfg is a hotbed of manufacturing research and development. CMfg's improved or entirely new manufacturing services are the result of ongoing improvements in IoT, SN, and virtual/augmented reality and simulation technologies.

CPPs, the most recent significant client need, also compel the maker to use these cutting-edge methods.

As a brand new paradigm with the potential to fundamentally alter manufacturing, the Internet of Things (IoT) has gained universal acceptance. Manufacturing equipment endowed with sensors and recognition and analysis and dissemination and actuator capabilities may be seamlessly integrated into a single system. It allows to new economic plus business potential used for the integral controller on this highly integrated smart cyber-physical realm.

3 U-Net Deep Learning-Based Hadoop Maintenance

In this section, a brief discussion of U-net-based Hadoop technique is implemented for future big data applications. Big data analysis is a set of information that may be used to provide accurate search results and relevant content for a more efficient work environment. Many semantic algorithms have been developed in the past to increase effective content searches; however, they have limits. Future apps will benefit from information retrieval and content filtering to make operations easier. The present state of online surfing and its recommendation system is that content tracking is erroneous, preventing users from obtaining the necessary information (Fig. 1).

In this work, an adaptive U-net-based semantic search strategy for advanced analytics systems has been proposed. Finally, performance metrics such as query time, building time, correctness, mean precise, stdError, and SSR are calculated. Here, the presented kvasir U-net architectural design interacts with existing systems and improves the accuracy and recall to 99.72% and 0.997%, respectively. This research improves the methods and can be computed with technology today.

$$H = \sum_{k=1}^{k} P_k \log P_k \tag{1}$$

$$G(\mathrm{dB}) = 2^* \log_{10}\left(Nt^{0.5}\right) = 10^* \log 10(Nt) \tag{2}$$

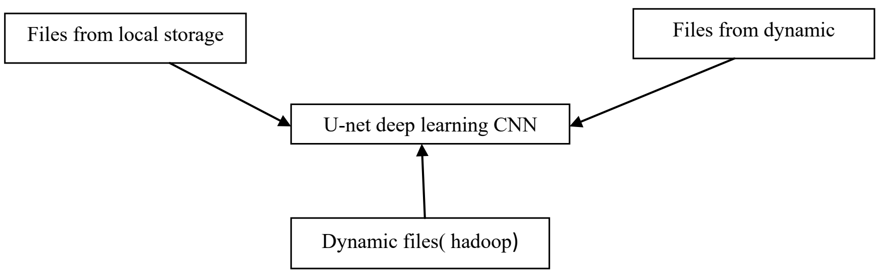


Fig. 1 U-net Hadoop system

$$X(m) = \sum_{n=0}^{N-1} x(n)[\cos(2\pi nm/N) - i\,\sin(2\pi nm/N)] = \sum_{n=0}^{N-1} x(n)\mathrm{e}^{-i2\pi nm/N} \quad (3)$$

$$\mathrm{CWT}_n(s, b) = \frac{1}{\sqrt{s}} \sum_{n=1}^{N-1} x(n)\Psi * \left(\frac{n-b}{s}\right) \quad (4)$$

The U-net described as the difficulty of banner $x(t)$ with wavelet limits ψa, $b(t)$, here ψa, $b(t)$ be enlarged and moved interpretation of wavelet work $\psi(t)$ and is portrayed as takes after:

$$\Psi_{a,b}(t) = \sqrt{a}.\Psi\left(\frac{t-b}{a}\right) \quad (5)$$

Autonomous parameters, i.e., a, b in this technique, is excessive and not capable of methodological implementations.

$$a_j = 2^{-j}, b_{j,k} = 2^{-j}.k \quad (j, k \text{ are integers}) \quad (6)$$

The disintegrating of the Hadoop into several unique recurrent packs down in U-net isolates. The high and low pass channels are utilized as a part of U-net that provides two courses of action: limits, scaling cutoff, $\Phi(t)$, and wavelet work, $\psi(t)$, separately.

$$\phi(t) = \sum_n h[n]\phi(2t - n) \quad (7)$$

$$\Psi(t) = \sum_n g[n]\Psi(2t - n) \quad (8)$$

On the other side, a wavelet work $\Psi_{j,k}(t)$ or scaling limit $\phi_{j,k}(t)$ that will be discretized at scale j and translation k might be procured from principal work $\psi(t)$ = $\psi 0{,}0(t)$ or $\Phi(t) = \Phi 0{,}0(t)$ by (Figs. 2, 3):

$$\phi_{j,k}(t) = 2^{-j/2}\phi\left(2^{-j}t - k\right) \quad (9)$$

$$A_1(t) = \sum_k a_j[k].\phi_{j,k}(t) \quad (10)$$

$$D_j(t) = \sum_k d_j[k].\Psi_{j,k}(t) \quad (11)$$

$$x(t) = A_j(t) + \sum_{j=-\infty}^{j} D_j(t) = \sum_k a_j(k).\phi_{j,k}(t) + \sum_{j=-\infty}^{j} \sum_k d_j[k].\Psi_{j,k}(t) \quad (12)$$

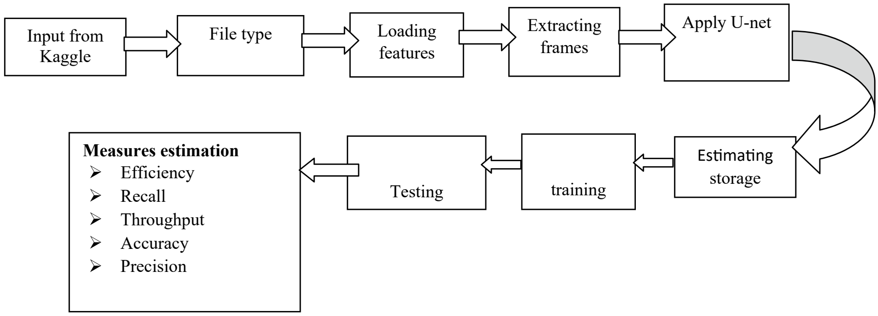
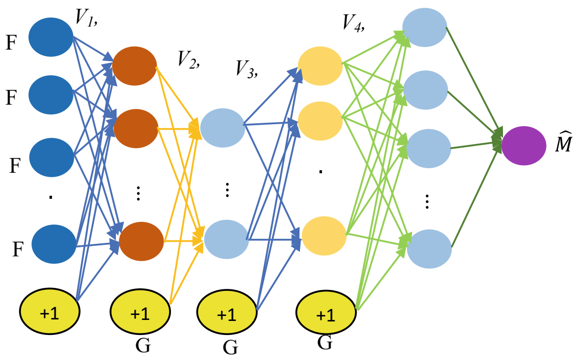


Fig. 2 Proposed block diagram

Fig. 3 Hadoop U-net deep stacked autoencoder



The Hadoop prognosis process is effectively performed by the deep-stacked autoencoder. The inputs have been sent through the activation function of the hidden layer to the cost function. The deep-stacked autoencoder learning, on the other hand, was done with the predicted U-net technique in order to improve the accuracy of classification results.

S. No	U-net-deep stacked autoencoder
1	Modify the population of the U-net layers
2	Define the coefficient matrix
3	Calculate the fitness for every investigate agent
4	Define the term $\frac{\rho}{t_\alpha}$, $\frac{\rho}{t_\beta}$, and $\frac{\rho}{t_\gamma}$
5	While $(m < H)$; H—maximum iterations
6	For every agent
7	Modify the location of an investigating agent
8	End for

(continued)

(continued)

9	Modernize the coefficient vectors
10	Calculate the fitness for all investigating agent
11	Modernize $\frac{\rho}{t_\alpha}$, $\frac{\rho}{t_\beta}$, and $\frac{\rho}{t_\gamma}$
12	$m = m + 1$
13	End while
14	Return to the best solution

The projected U-net technique delivers improved classifier performance in Hadoop file management by combining the layer sampling feature and the imitating model of atomic search. To improve the classification results, the interaction force between the files is added to the search local server position update equation. The proposed U-net technique is based on the layer investigation performance and the Hadoop advanced analytics imitating paradigm.

Industrial big data will be generated as a result of the deployment of big data with ubiquitous sensing capabilities in production, changing physical things and operators into "cyber-ones." The big data and Internet of Things (IoT) may be utilized to help manufacturers make the most use of cloud-based dynamic services. Big data have never been properly used before, even though it contain valuable information and knowledge that may help manufacturers make better decisions. Analysis of the huge IoT data from design to potential and outstanding concerns were examined. Industrial IoT's success or failure depends on big data, which is meant to provide businesses with valuable information.

4 Results and Discussion

In this section, a brief investigation is performed, a practical methodology called U-net deep intelligence autoencoder and ENR classifier is proposed to Hadoop big data maintenance. At pre-processing stage, the files of Kaggle dataset is processed to the U-net block. In this stage, the files have been trained, and the unstructuring is eliminated. The unwanted files are removed by using the primary step of the U-net method. After this stage, the U-net regression machine learning algorithm is applied to processed data. Inflexible net regression, seizure classification, and pre-diagnosis are implemented based on a balanced weight system. The encoder used in this work is a U-net autoencoding scheme; this facility is used to reveal a better file accessing system. Finally, the U-net model achieves improved performance metrics, i.e., accuracy = 98%, specificity is 93%, and sensitivity is 98%. It is better achievement than current methods (Table 1).

Big Data Analytics and Cooperative Processing of Large-Scale Data. The results of various levels and phases of factors provide real-time data for various goals

Table 1 Comparison of results

Models		NB + KNN	Nonlinear multi-domain	Deep stacked	AWGO deep stacked	U-net model
Training data	Accuracy	87.9623	91.9692	92.6	94.1035	98.78
	Specificity	88.8321	91.243	91.8	91.9832	92.32
	Sensitivity	58.6483	84.41	97.22	97.3456	98.52
K-fold data	Accuracy	92.1374	93.26	93.45	93.6431	97.732
	Specificity	91.2389	91.56	91.8	91.984	95.74
	Sensitivity	92.6552	92.8	92.3	93	96.94

such as quality control and energy usage. Machine conditions and job progress are other examples of these. Data aggregation across industrial systems, both vertically and horizontally, is a problem that has to be solved. For retrospective review and future prediction, data are gathered in several dimensions throughout the timeline. To decrease the quantity of network traffic, cleaning and aggregating data at the devices' ends may be an option. This might be a never-ending process of improvement. It is necessary to have a collaborative data processing approach that is flexible and evolutionary between local nodes and the cloud.

Figure 4 clearly explains about various technology outcomes in this proposed model that attain more improvement (Fig. 5 and Table 2).

Table 3 clearly explains about various datasets generating measures comparison. This Hadoop system gives accurate measures like accuracy 87.46, time 0.724 s, and storage space of 192 bytes (Fig. 6).

Mutual information-based individual feature extraction on Hadoop data, this algorithm attains 56.9% means for selecting the Kaggle dataset.

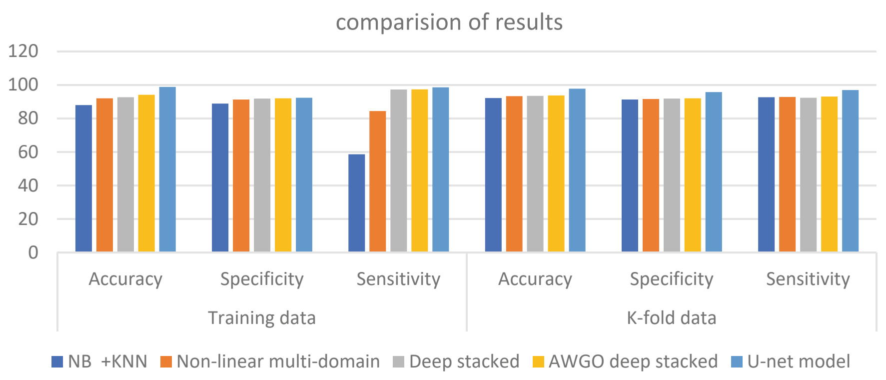


Fig. 4 Comparisons of outcomes

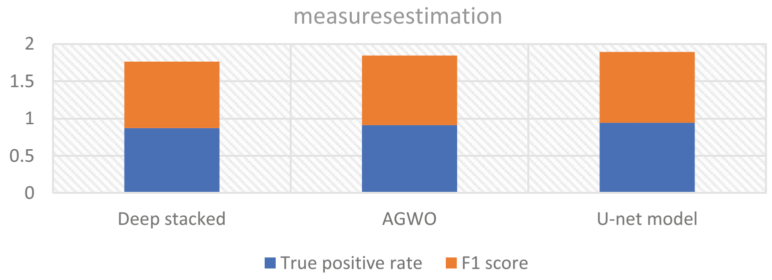


Fig. 5 U-net analysis

Table 2 Measures estimation

Parameter	Deep stacked	AGWO	U-net model
True positive rate	0.872	0.912	0.943
F1 score	0.893	0.934	0.951
MNSE	0.062	0.04	0.01

Table 3 Comparison of datasets

Method	Accuracy in %	Time in s	Storage bytes
Kaggle	85.23	1.971	259
spark	86.71	1.329	187
Hadoop	87.46	0.724	192

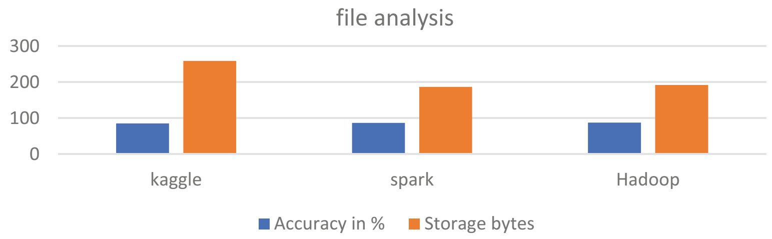


Fig. 6 Comparisons of datasets

5 Conclusion

Many applications, such as cloud, server, and Hadoop maintenance, can be given by an advanced big data platform. The current big data infrastructure, which uses Hadoop modelling, is unable to provide dynamic actions in a timely manner. The application's robustness is being harmed by latency and storage concerns. The Internet and big data platforms have also been underutilized in cloud storage upkeep to drive the manufacturing process. Continuous development and evolution of big data cloud systems will be aided by large data handled at cloud gateways. This research additionally shows the most recent way (DL-enabled operations Facilities) via a sponsoring software arrangement, which may encourage swift procedure and updating the large data file maintenance procedures in a closed loop with intelligent video surveillance. This research could aid cloud-based large data file production systems in operating more rapidly and efficiently. Big data analysis is undertaken in this research project, U-net, which is built on Hadoop and Sparks. On the python 3.7 software tool, the following application is verified. Accuracy of 98.78%, sensitivity of 98.52%, and throughput of 99.23% were achieved, all of which beat the approach. This big data analytics solution based on the U-net competes with current technology.

References

1. Yang C, Lan S, Shen W, Huang GQ, Wang X, Lin T (2017) Towards product customization and personalization in IoT-enabled cloud manufacturing. Clust Comput 20(2):1717–1730
2. Yang C, Shen W, Wang X (2018) The internet of things in manufacturing: key issues and potential applications. IEEE Syst Man Cybern Mag 4(1):6–15
3. Shi Y, Han Q, Shen W, Zhang H (2019) Potential applications of 5G communication technologies in collaborative intelligent manufacturing. IET Collaborative Intell Manuf 1(4):109–116
4. Yang C, Lan S, Shen W, Huang GQ, Wang L (2019) Software-defined cloud manufacturing in the context of industry 4.0. In: 2019 WRC symposium on advanced robotics and automation (WRC SARA). IEEE, pp 184–190

Efficient Model for Predicting Cyber Hacking Breaches

Kothandaraman D, K. Rajchandar, Bhavana Jamalpur, M. Nikhilvardhan, and Bura Vijay Kumar

Abstract The cyber occurrence informational indexes as a means of expanding our understanding of the current threat situation is essential. More research is needed on this topic before finishing some research projects. This paper compares a break-in information collection to 12 years (2005–2017) of cyber hack exercises required malware invasion, and the results are quantifiable. Since autocorrelations are shown by both hackers break occurrence and penetration size, it shows that stochastic processes rather than conveyances should be used to depict these data. Stochastic method models are proposed to fit the break sizes and between appearance times. The informational index is being subjected to both subjective and quantitative pattern inspections in order to gather more information about the progression of hacker penetration occurrences. Despite the fact that cyberattacks are becoming more frequent, they are not getting much worse in terms of the damage they can cause.

Keywords Cyberattack · Stochastic process · Hackers

1 Introduction

An information breach is a security incident in which sensitive, guaranteed, or secret data are accessed by an individual who was not authorized to do so. A data breach is the intended or accidental disclosure of private or confidential information to an untrusted environment. Coincidental information disclosure, data leakage, and other forms of data leakage are all possible explanations for this phenomenon. There are a variety of ways that this could happen, including a break-in or the stealing or drop of advanced technologies, such as tape drives, disk drives, or mobile devices, which contain these type of sensitive data, posting it online or on a PC that is widely reachable via the Internet without adequate information reliability safeguards, or

Kothandaraman D (✉) · K. Rajchandar · B. Jamalpur · M. Nikhilvardhan · B. Vijay Kumar
School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India
e-mail: dramanko@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_42

trading it with an establishment that is not completely open but is not officially approved for secrecy. Data breaches are still a major problem, despite the use of mechanical game plans to protect complex systems. This prompts us to depict the evolution of data bursts. As a result of this, we will have a better understanding of data breaches, as well as a better understanding of the many systems in place to prevent them, such as security [1]. Despite the fact that many people believe that insurance is vital, the existing understanding of data breaches puts a limit on the advancement of accurate cyber risk predictions that may be used to manage insurance costs. In this paper, we demonstrate how obligations can be carried out in practice [2].

In this paper, we argue that instead of circling the cracks, we should use a stochastic technique to show both the burst size and burst passage timings. In this paper, we appear that continuous plan models can predict the time between landing and the burst size precisely. These computerized danger factors should be shown using stochastic techniques rather than disseminations, as far as we can tell from this paper. We show that a certain copula can adequately represent the link uniting the location entering time and the break measurements [3]. In this paper, the closeness of this reliance is demonstrated, as well as the consequences of ignoring it. It is also crucial to consider the reliance when predicting bury section times and break measures, as the findings are not always perfect. Our hopes are that the current assessment will lead to more examinations that will deliver significant experiences in exchange for the nearing of relief. Insurers, government agencies, and controllers will find this information useful because they must have a thorough understanding of the concept of information permeating hazards. We hope that the current investigation will lead to more studies that can provide valuable insights into exchange risk reduction. Insurance companies, government agencies, and controllers can benefit greatly from this kind of information, as they must have a thorough understanding of the threats associated with information penetration [4].

2 Related Work

Earlier works enthusiastically associated with the current review [5] inspected a dataset [6] of 958-man nature incident events passed off the US between the years 2000 and 2008. They found that the man or woman persona misfortunes as per episode, appeared through X, can be shown by an amazing tail stream Qs(X > m)—m − α where α = zero.7 ± zero.1. This result visits standard size while separates the dataset unsurprising with sort of workplaces: business, practice, authority, and objective the situation. For the clarification that possibility width limit of the person setbacks in a state of harmony with event is static, the circumstance of man or woman catastrophe is consistent from mark of the break measurement.

Isolated an exchange break dataset [7] of two, 253 break episodes that extent lengthy than 11 years (2005–2016). Those break episodes integrate two orders: unwise breaks (i.e., events added about by stray, inclined of, taken contraptions, or prohibitive reason) and noxious infiltrating (i.e., episodes added around through hacking, insider and specific expectations).

The progressive break events dataset [8, 9] runs extended more than 11 years (years 2000–2016). They applied ridiculous worth thought [10] to review the main outrageous phenomenal break measurement, and similarly supported the huge break measurements by means of a twice condense Pare-to dispersing. They also applied directly lose the faith to ponder the repeat of the information breaks, and affirmed that the repeat of sizable entering events is free of time for the USA associations, but appears an enlarge plan for non-USA associations.

The dependency among advanced dangers of multi levels is as follows: inner an affiliation (inside dependence) and every through office (in general reliance). The applied the Archimedes copula to show computerized perils accomplished by means of sickness events and affirmed that there exists a little dependency between those dangers. Applied a copula fundamentally builds probabilistic knowledge neighborhood overview advanced mark of concern. Explore to use copulas to show junior computerized chances [11]. Applied copulas to find dependency expert while showing the common sense of advanced confirmation primary-alerted [12].

Assessed bivariate organization insurance bets with dependency. Separated and each one in all that about assessments alluded to over, the contemporary is one in every each of the a benevolent it uses more machine to isolate one more view of perspective of break episodes (computerized hacking break events) [9]

The broke down the place of computerized risks by using a dataset amassed at neighborhood. Use dataset accumulated at desirable, misused the legitimate homes which consolidates expanded run reliance and noteworthy characteristics to painting and really rely on how much goes after toward the honeypot; a regularity assessment of a linked dataset is represent in [13].

There is a correlation between the idea of framework breaches and attacks on framework activity and functioning. In order for a framework to fall apart, it can be subjected to dynamic or uninvolved assaults. In the event of an attack on a system, the data security is breached, allowing the programmer to hack or steal all of the system's data [14].

The past assaults on the framework, as well as similar assaults around the earth and the specific representations, are taken into account when decide the possibility of future assaults. Insuring the framework against intrusion is a top priority for the security and insurance teams. Even though many remedies have been implemented by the framework head, the cyberattacks continue to take place, and the risks and threats they pose are evaluated. In the event that countermeasures are employed during the cyberattack, the penetration likelihood can be processed by the general break likelihood [15].

3 Proposed Method

The present investigation is based on a number of unanswered questions, such as is the number of cyberattack-related data breaches increasing, decreasing, or stabilizing? We can gain a better understanding of cyber risks if we respond to this issue in a principled manner.

When asked this question in the past, no answers were given. Particularly in [7], the dataset only covered the years from 2000 to 2008 and did not contain any breaches created by cyber-assaults; the dataset examined in [8], on the other hand, is more current and consists of varieties of incidents: careless breaches (i.e., breaches resulting from lost, disposed of, taken devices, and other motives) in addition to malevolent breaches.

We don't include careless breaches in our current review because they are more likely to be the result of human error than a cyberattack. In order to focus on the hacking subcategory (hereafter referred to as the hacking breach dataset from that point on), the paper examine the insider, installment card misrepresentation, and obscure breaches contained in [8], but also note that each of the other three subcategories are fascinating in their own right and should be researched independently. We display the both hacking breach event inter-arrival times and breach measurements, which analyzes the type of hacking attacks.

Advantages of proposed system are to understand the uniting arrival times, the break measurements, and attack. Both individual and quantifiable pattern inspection of the digital hacking break episodes are shown in Fig. 1.

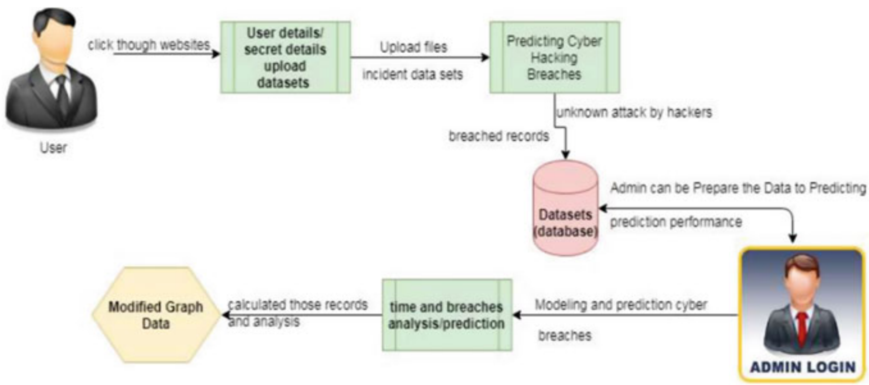


Fig. 1 Working model of proposed system

3.1 Upload Data

Both administrators and authorized users can upload data resources to the database. It is possible to submit data using a key in order to ensure that the user's private information is not leaked without their permission. A user's details are shared with the administrator, and the administrator can grant or deny access to each user. Access to the system is restricted to those who have not been granted permission to do so.

3.2 Access Details

Data in a database can be made available only to those authorized by the database's administrators. It is the administrator's responsibility to keep track of all uploaded data, and the authority to accept or disapprove users depending on the information they supply. Administrators can provide users access to data in the database. Admin manages the uploaded data, and admin is the only one who has the authority to process the accessed information.

3.3 User Permissions

Access to the data from any source is granted only with the administrator's approval. Admins give users permission to share and verify their personal information before granting them access to their data. If a user tries to access the data in an incorrect manner, they will be blocked. If a user asks the administrator to unblock them, the administrator does so based on previous requests and activity.

3.4 Data Analysis

Graphs are useful tools for data analysis. In order to acquire the most accurate scanning and forecast of the dataset and the supplied data regulations, the data collected are applied to a graph. There are numerous ways in which the data can be evaluated using this visual representation. Graphs are used to perform data analysis. To acquire the better investigation and prediction of the dataset and specified data regulations, the collected data are applied to a graph. This pictorial representation can be to study the dataset and have a better understanding of the data contents.

3.5 Algorithm

Calculation Utilized: Calculation for Prediction.

The VzRα's of the Hacking Occurrences Between Appearance Times and the Break dimensions individually.

Input: Verifiable episodes between appearance times and break measurements, implied by {(dtj,ytj)}j = 1,…,p + q, where an in-model {(dtj, ytj)}j = 1,…,p as referred to more was using for placing and an out-of test {(dtj, ytj)} i = p + 1,…,q is use for appraisal figure precision; α level.
Step 1. Start.
Step 2. for j = p + 1,…, q.
Step 3. Measure the LACB1 representation of the events between appearance considering {dc |c = 1,..., j − 1}, foresee the contingent mean Ψj = expr (ω + a1 log(ψj − 1) + b1 log(ψj − 1)).
Step 4. Measure the ARMAA-GARCHH model of changed dimension, and anticipate to accompanying mean μj quality mix-up σj.
Step 5. Choose an acceptable Copula used the multivariate remaining from the past models taking into account.
Step 6. Considering the evaluated copula and reenact 10,000 two-layered copula tests.
Step 7. For the events between appearance times and convert the reenacted subordinate models v(h) 1,j's to the N(h) 1,j's use the opposite of the evaluated summarized gama scattering, h = 1,…, 10,000.
Step 8. For the break sizes and convert the copied subordinate models v(h) 2,j's into the N(h) 2,j's use the opposite of the surveyed varied ridiculous worth transport, h = 1,…, 10,000.
Step 9. Process the expected 10,000 two-layered break information.
Step 10. Process the VzRα,d(j) for the episodes between appearance times and VzRα,y(j) for changed break measurements taking into account the reenacted break information.
Step 11. if b(h)j > VzRα,b (j).
Step 12. An encroachment to the events between appearance time happens.
Step 13. end if.
Step 14. if y(h) J > VzRα,y(j).
Step 15. An encroachment the break measure happens;
Step 16. end if.
Step 17. end for.

The circumstance of digital hacking breaks reflects the final product of the digital attack obstruction collaborations (e.g., regardless of whether the attack devices can accurately avoid the safeguard contraptions). Yet again notwithstanding the truth that the particular wonder referred to above can happen under a broad assortment of conditions and effectively making sure about its inspiration is past the amount of the contemporary (basically because of the shortfall of different sorts of aiding records), a single opportunity is the going with; while the assault contraptions are not the slightest bit convincing escape the assaults point, the aggressors could likewise need to separate a longer season of endeavor to develop new attack gadgets for practically breaking measurements.

4 Results and Analysis

In the table, all the files have been loaded in the previous table, now click the add data button to load additional files and obtain results as shown in Table 1.

Table 2 is a malware analysis is the procedure of figuring out how a doubtful file or URL acts and what it is trying to accomplish.

Table 3 unmalware data are the process of files, and URL is used for getting good results.

Table 1 Add data

Type	Method
Entity	Accendo Insurance Co,
Year	2015
Records	23,344,232
Organization type	Health
Methods	Cyber attack
Add data	Website Url
Time	12:00 pm

Table 2 Malware data

Entity	Year	Records	Type	Method
21st Century Oncology	2014	2,300,000	Health care	Hacked
Accendo Insurance	2012	185,350	Health care	Poor security
Adobe Systems	2014	162,000,000	Tech	Hacked
Ameritrade	2006	300,000	Financial	Lost/stolen media

Table 3 Unmalware data

Data	Entity	Attack results
Website Url1	Accendo Insurance Co	Unmalware data
Website Url2	21stcentury Oncology	Unmalware data
Website Url3	Adobe Systems	Unmalware data
Website Url4	Ameritrade	Unmalware data

Table 4 Breaches analysis

Malware data	Network position	Method
Man-in-the-middle attack	Hacked	4
Drive-by-attack	Hacked	36
Password attack	Lost/stolen media	10
Eaves dropping attack	Lost/stolen media	8
Sql injection attack	Hacked	34

Table 4 shows breach analysis information is stolen or taken from a system without authorization.

The efficient model for cyberattacks to predicting methods login screen user sign up details entered and then click on register and upload data to get results.

Figure 2 shows malware breaches analysis in year wise predicted results that represented using line graph and Fig. 3 shows breaches analysis using bar graph.

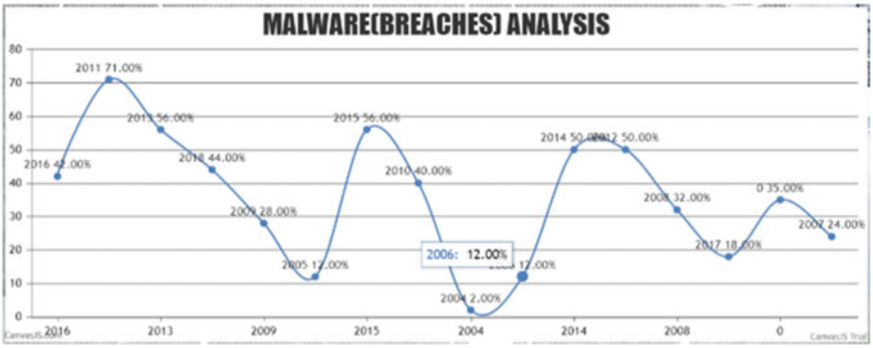


Fig. 2 Breaches analysis with line graph

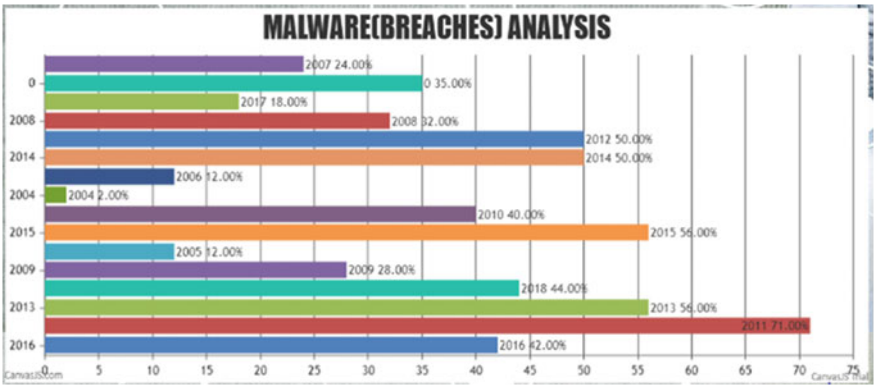


Fig. 3 Breaches analysis with bar graph

5 Conclusion

The paper looked at a hacking breach dataset and discovered that instead of conveyances, the inter-arrival length and breach magnitude must be displayed using continuous processes. Models with strong fitting and expectation accuracy are now available. To expect the cumulative possibility of an event with a particular breach range occurring in the future recommend utilizing a copula-based technique. The proposed systems surpass those proposed in the writing because the latter forget temporal connections as well as the relationship uniting event inter-arrival intervals and breach measurements. To learn more subjective and quantitative research. The paper learned a lot about cyber security, including the fact that cyber hacking breaches are becoming less common but not less serious. The method presented here can be used to similar datasets. Many questions remain to be answered. Analyzing approaches to foresee significant attributes and handle lacking facts, for example, is both enjoyable and demanding (i.e., breaking occurrences are not described). Examining when breaches occur is also crucial. Finally, greater research into breach consistency is required (i.e., upper level of assumption accuracy).

References

1. Mailart T, Sornete D (2010) Heavy-tailed distribution of cyber-risks. Eur Phys J 75(3):357–364
2. Edwaards B, Hofmeyr S, Forrest S (2016) Hype and heavy tails: a closer look at data breaches. J Cybersecur 2(1):3–14
3. Wheatley S, Mailart T, Sornete D (2016) The extreme risk of personal data breaches and the erosion of privacy. Eur Phys J B 89(1):1–7
4. Herath H, Herath T (2011) Copula-based actuarial model for pricing cyber-insurance policies. Insur Markets Companies Analyses Actuarial Computations 2(1):7–20
5. Mukhopadhayay A, Chatterjee S, Saha D, Mahanti A, Sadhukhan SK (2013) Cyber-risk decision models: to insure it or not? Decis Support Syst 56:11–26
6. Xuu M, Huaa L, Xuu S (2017) A vine copula model for predicting the effectiveness of cyber defense early-warning. Technometrics 59(4):508–520
7. Xuu M, Huaa L (2017) Cybersecurity insurance: modeling and pricing
8. Pengg C, Xu M, Xu S, Hu T (2018) Modeling multivariate cybersecurity risks. J Appl Statist 1–23
9. Eling M, Loperfiido N (2017) Data breaches: goodness of fit, pricing, and risk measurement. Insur Math Econ 75:126–136
10. Bagchii KK, Udo G (2003) An analysis of the growth of computer and Internet security breaches. Commun Assoc Inf Syst 12(1):46
11. Condon E, He A, Cukier M (2008) Analysis of computer security incident data using time series models. In: 2008 19th international symposium on software reliability engineering (ISSRE), pp 77–86
12. Praveen P, Shaik MA, Kumar TS, Choudhury T (2021) Smart farming: securing farmers using block chain technology and IOT, EAI. Springer Innovations in Communication and Computing, pp 225–238
13. Mothe R, Balija S, Chanti Y, Kumar BV (2018) A modified fault diagnosis scheme in wireless sensor networks. Int J Eng Technol 7(8):226–229

14. Kothandaraman D, Amuthan A, Chellappan C, Sreenath N (2010) Prevention of wormhole and rushing attack on location-based geocasting and forwarding routing protocol in mobile ad-hoc network. Int J Comput Appl 957:8887
15. Böhmme R, Kataaria G (2006) Models and measures for correlation in cyber-insurance. In: Proceedings of workshop on the economics of information security (WEIS), pp 1–26

Detection of Cyberbullying in Social Media Using Machine Learning Techniques

S. Arvind, M. Sriram, R. Reshma, P. Vaishnavi, and S. V. Devika

Abstract Cyberbullying is the act of threatening someone while using technology. Social networking networks provide a thriving platform for threats who use them to harass or harm defenseless young adults. We can identify language patterns used by threats through machine learning, and we can develop criteria to recognize online harassment as a result. The majority of research on machine learning-based cyberbullying detection has been conducted on languages including English, Chinese, and Arabic. Regional Indian languages have not been the subject of many studies. We have created a methodology in this research that can identify cyberbullying content in a rare or rather regional Indian language.

Keywords Cyberbullying · Machine learning · Random forest · Passive aggressive classifiers

1 Introduction

Although very different from traditional harassment, cyberbullying is nevertheless torturous. The effects and dangers are still present, if not more severe and widespread. Even though abuse occurs on websites rather than in person, cyberbullying should still be treated seriously. When it is appropriate, cyberbullying can take many different forms. It does not actually entail altering someone's profile or posing as someone else. It also includes circulating rumors to disparage someone or making disparaging things about them online. Bullying via social media or online Bullying includes actions and strategies designed to oppress, irritate, or stigmatize any person. These heinous acts significantly affect everyone and have a powerful and negative impact on them. Basically, they take place through online media and public gatherings and other online sites.

S. Arvind · M. Sriram · R. Reshma (✉) · P. Vaishnavi · S. V. Devika
Hyderabad Institute of Technology and Management, Hyderabad, India
e-mail: scarvi@rediffmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_43

2 Purpose

Cyberbullying has increased dramatically over the past ten years, and it does not just happen in English; it also happens in other languages. Because cyberbullies thrive in large crowds, it is crucial to recognize cyberbullying in varied linguistic contexts. As a result, the proposed model that we are going to create will aid in identifying cyberbullying content in Bengali, one of India's more vocal regional languages. In this language, very few studies on cyberbullying have been conducted. The cyberbullying models utilized in earlier English-language studies will be the ones we employ here. Despite the possibility that execution and performance may vary due to semantic differences between English and non-English contents, to combat such problems, this model proposes the utilization of machine learning algorithms and the consideration of user data for detecting digital harassing on text.

3 Existing System

From datasets from four websites, Hsien [1] employed a method combining keyword matching, opinion mining, and social network analysis and obtained a precision of 0.79 and recall of 0.71. A troll (someone who engages in cyberbullying) on a social networking site may always have a real profile to see how other users perceive the phony profile, according to a theory put forth by Patxi Gal'an-Garc'a et al. [2]. They suggested using machine learning to find these profiles. The identification method looked at a few profiles that are somewhat related to them.

From datasets from four websites, Hsien [1] employed a method combining keyword matching, opinion mining, and social network analysis and obtained a precision of 0.79 and recall of 0.71. A troll (someone who engages in cyberbullying) on a social networking site may always have a real profile to see how other users perceive the phony profile, according to a theory put forth by Patxi Gal'an-Garc'a et al. [2]. They suggested using machine learning to find these profiles. The identification method looked at a few profiles that are somewhat related to them.

For instance, a trolling account might not have an actual account to trick such systems, or there might be specialists who can alter writing patterns and behavior. More effective algorithms will be required to adapt to shifting writing styles.

A collaborative detection technique was proposed by Mangaonkar et al. [3] in which data and outcomes are merged to provide results from many connected detection nodes that utilize either the same or different algorithms. A concentration-based B-LSTM method was proposed by Zhou et al. [4]. KNN and fresh embeddings were employed by Banerjee et al. [5] to achieve a precision of 93%.

4 Areas of Concern

A vocabulary is not created from every document. The vocabulary could be made up of all words (tokens) used in all papers or just a few of the more frequent ones.

The TFIDF technique uses the same procedure to generate a vocabulary in order to obtain its features, which sets it apart from the bag-of-words model.

5 Literature Survey

In the paper [1], the authors organized a dataset using a Java program designed to separate Bangla Text discussions from online media platforms, specifically Facebook and Twitter. In addition to the discussions, they also collected client segment information from Twitter using Twitter Rest Api, which was marked physically and combined with sack-of-words approach for the model. They created a model using machine learning techniques like SVM, KNN, Naive Bayes, and J48, then assessed how well the various techniques were explained. They conducted the tests in two stages, first preparing and testing the model using content debates in Bangla Text, and then included both content-based highlights and user information in the second stage.

SVM outperformed a wide variety of calculations in the two stages. Paper [2], one of the few publications we could find that attempted to identify harmful Bangla Text, presented unigram string highlights to enhance results in addition to using a root level algorithm to identify oppressive content. They analyzed the experiment with a variety of different string properties, including unigram, bigram, and trigram, to find out with what kind of highlight their proposed calculation performed better. The unigram features do not take into account the word importance in a single sentence.

However, using this component tends to reveal words that were more destructive. They constructed a dataset of text conversions and comments from the comment sections of certain well-known Facebook sites' public postings for the paper [3]. The dataset was separated into two groups, bully and non-bully, with the bully group further subdivided into four groups: sexual, troll, religious, and threat. The gender of the person who wrote the comment and the person it was addressed to were also added to the dataset. The comments in the dataset also include additional user-specific information like occupation and, finally, the total number of responses to a given comment.

They provided analyses for each column in the dataset in order to offer more information about the dataset. The availability and volume of information in each category provide a good foundation for building a machine learning model that can distinguish various forms of cyberbullying in Bangla. Machine learning can be used to identify linguistic examples of tyrannical jerks, and it can also be used to construct a model to recognize digital tormenting acts. In this way, the central goal of paper [4] was to present a controlled machine learning model for identifying and preventing digital harassment in English. A dataset from Kaggle that contained content related to online abuse was used to evaluate the model.

On both TFIDF and sentiment analysis extraction approaches, the effects of SVM and neural network classifiers were considered. It was determined that a neural network outperformed an SVM classifier in terms of performance. Additionally, this work was compared to a related study that used the same dataset, finding that neural networks outperformed their classifiers in terms of precision and F-score. They established their own API and scraper to collect the dataset from various social media platforms, newspaper reviews, and tour evaluations. They proposed a model in the paper [5] that offers a multilingual cyberbullying detection strategy in multiple Indian languages, primarily Hindi and Marathi. The dataset was loaded into the machine learning algorithm after hand labeling and removal of any redundant text, stop words, and special characters' models.

The model was created using machine learning languages like Multinomial Naive Bayes, Stochastic Gradient Descent, and Logistic Regression. They categorized the information using the "bag of words" method, which ignores the syntax as well as the order or appearance of the terms while retaining their frequency. They assessed the F1-scores of each method to gauge accuracy and discovered that LR surpassed all other ML techniques with the lowest error rate. They provided a summary of multilingual cyberbullying recognition in paper [6]. After conducting extensive study, they discovered that the majority of work detecting cyberbullying has been done primarily in English, and thus, they tried to identify cyberbullying in order to give their model a unique characteristic in Arabic language. In their research, they dealt with detecting cyberbullying using a variety of ML techniques.

They collected a dataset of 32 thousand tweets, of which 1800 were classified as harassing messages. With the aid of algorithms like Support Vector Machine (SVM) and Naive Bayes, they were able to accurately identify cyberbullying with an accuracy of roughly 92 and 90%, respectively. When compared to earlier models for English bullying detection, this framework's results were not entirely accurate. However, the purpose of this work was to illustrate that cyberbullying in Arabic is quite recognizable, showing us that it is entirely possible to spot harassment in other regional or unique languages as well.

6 Proposed System

This project's solution to the challenge of detecting cyberbullying involves categorizing content as either containing or not including the two main types of cyberbullying: hate speech on Twitter and personal attacks on Wikipedia.

Tokenization: Tokenization is the process of dividing unprocessed text into meaningful words or tokens. For instance, the phrase "we will do it" can be tokenized as "we," "will," "do," and "it." Sentence tokenization and word tokenization are two different types of tokenization. Although there are many other types of tokenization, we utilize the Regex Tokenizer for this project. A regular expression is used in the Regex Tokenizer to determine the tokens to be used. The following regular expression is used to select tokens: e.g., all of the alphanumeric tokens are extracted for the regular expression "w+ ."

Stemming: Stemming is the process of changing a word into a root word or stem. For instance, the stem for the phrases "eating," "eats," and "eaten" is "eat." Since the root word "eat" has three branches, all three of them should be understood to mean the same thing. Porter, Lancaster, Snowball, and Regexp stemmers are the four varieties of stemmers that NLTK offers. The subsequent project makes use of PorterStemmer.

Stop word Removal: Stop words are words that offer no additional sense to a sentence, such as the terms what, is, at, and a. You can delete these words because they are unnecessary. A list of English stop words is included in NLTK and can be used to filter out all tweets. When we train deep learning and machine learning models, stop words are frequently eliminated from the text data since the information they provide is unrelated to the model and aids in enhancing performance.

7 Implementation

A. Training and Testing Dataset: In this module, we will narrow our attention to the dataset we have acquired, initially removing all rows with null entries. At that time, we will remove any extraneous features that can jeopardize the correctness of our algorithm. The dataset will likewise be split into training and testing sections in this instance. We will use 80% of the dataset to train the model, and the remaining 20% will be used to assess the accuracy of the training model. The collected information is manually classified as bully (sexual, threat, troll, or religious) or not-bully. The dataset also includes three additional variables that identify the gender of the commenter, the category of the remarks passed, and the overall number of responses for each comment.

B. Data Preprocessing: Since there were signs of unstructured material in the data that were obtained, preprocessing was necessary. It is essentially meant that in order to increase accuracy, we needed to clean or trim the data. The preprocessing of the data required a number of stages, including data cleansing, stop word removal, and tokenization. We removed any extraneous terms from every text exchange in accordance with the Bengali lexicon with the aid of a stop word filter. Stop words are those words that do not provide any information that can be used to determine what category a text belongs in. We converted all of the data to lowercase letters to facilitate the subsequent processes with the goal of eliminating the distinction between capital letters and lowercase letters.

Furthermore tokenization had to be practiced on these text contents to facilitate the feature extraction step. Tokenization can be defined as a way of separating or isolating every word that compiles in a document or even a conversation.

C. Feature Extraction: The text conversations from the preprocessed data will be turned into a vector space model, where they will be described by a vector of features that were retrieved using Term Frequency Inverse Document Frequency (TFIDF). TFIDF is mostly used to gauge or assess how pertinent a term is to a document or group of documents. The primary feature of TFIDF is that it effectively analyzes text and determines the weights of these words in relation to the document or sentence. We will also employ word-level feature extraction techniques in addition to TFIDF; this particular technique is referred to as "Bag of Words" or "Bag of n-grams" representation.

While fully ignoring the position or sequence of the words in the text, it implies that documents are defined or represented by instances of the words. A common parameter for combining the vectorizer and a machine learning model is max df, which is used to eliminate phrases from the document that appear too frequently.

D. Classification: The collected characteristics are then used in the proposed model's classification phase, which entails training and testing the classifier to see if it can correctly identify cyberbullying. The Support Vector Machine (SVM), Logistic Regression (LR), Random Forest, and Passive Aggressive (PR) classifier are just a few of the machine learning techniques and algorithms we will employ. Several assessment lattices are used to evaluate each of these classifiers. Accuracy, precision, recall, and F-score are some of these requirements.

8 Result

The graph shown in Fig. 1 represents the percentage of offensive and non-offensive tweets entered in our system. As per the graph, it is observed that there are 64.44% of non-offensive content and 35.56% of offensive content from the data entered.

The pie chart shown in Fig. 2 represents the percentage of offensive and non-offensive data entered in our system. From the chart, it can be observed that the majority of the data is non-offensive.

The bar graph shown in Fig. 3 represents the accuracy percentage of different algorithms used in our model. From the graph, it can be observed that Support Vector Machine algorithm shows the most accurate result and Naïve Bayes algorithm shows the least accurate result.

Figure 4 shows the line graph representing the accuracy of the algorithms used in our model. From the graph, it can be observed that Support Vector Machine algorithm shows the most accurate result and Naïve Bayes algorithm shows the least accurate result.

RESULT

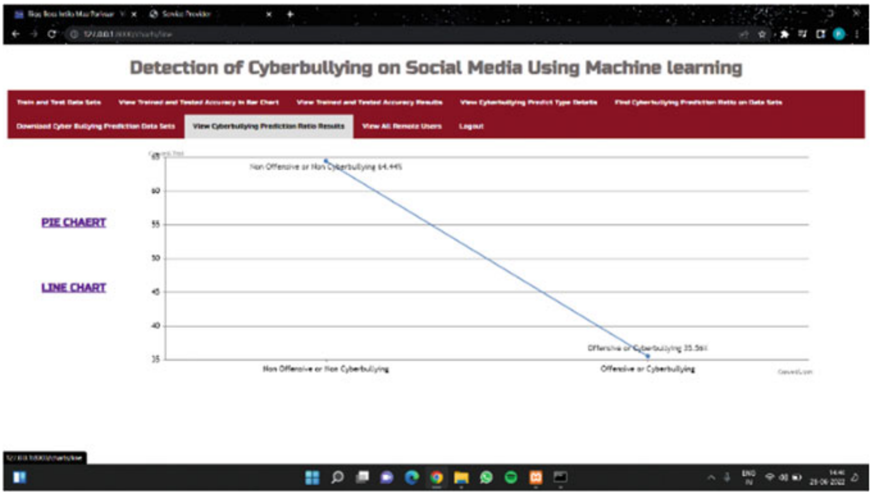


Fig. 1 Line graph showing percentage of offensive and non-offensive data

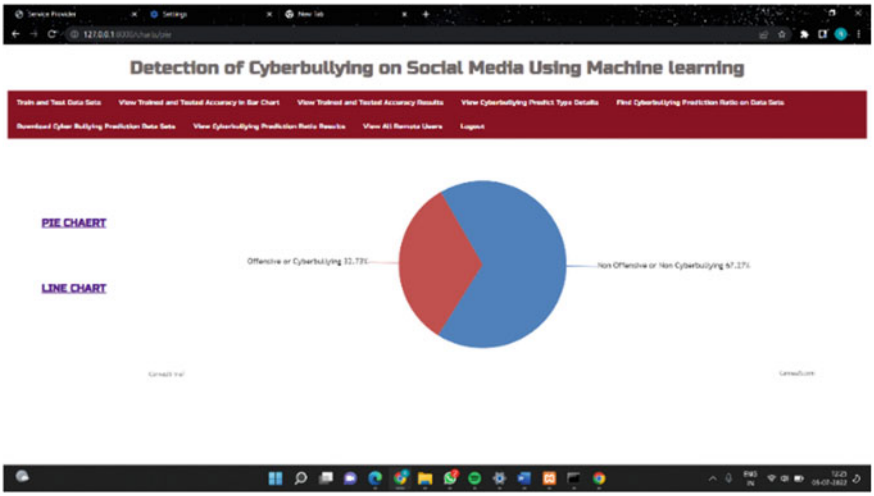


Fig. 2 Pie chart showing percentage of offensive and non-offensive data

9 Conclusion

We attempted to use text classification techniques to identify cyberbullying in Bengali. Even though we have used a variety of text-based classification algorithms, including SVM, Logistic Regression, Passive Aggressive, and Random Forest, other

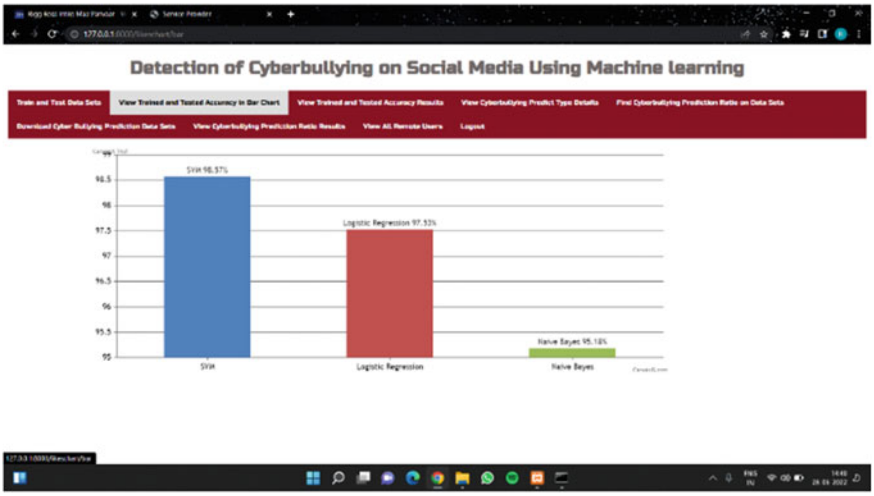


Fig. 3 Bar graph showing accuracy of algorithms

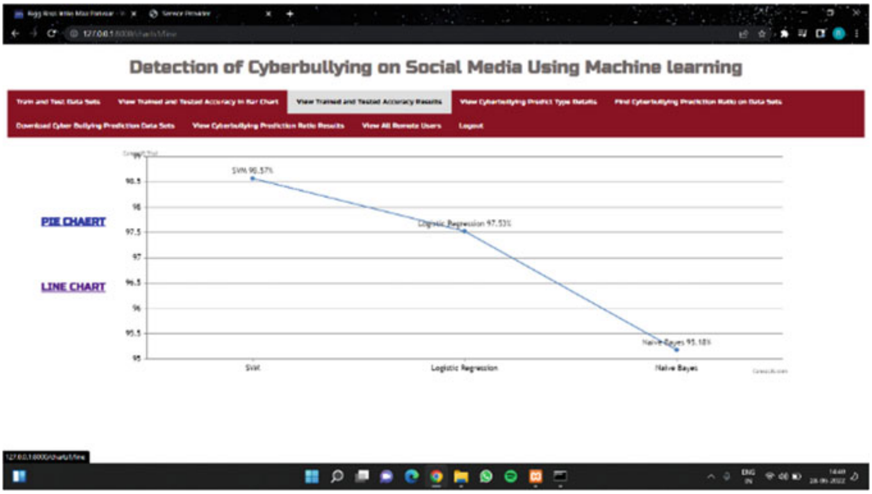


Fig. 4 Line graph showing accuracy of algorithms

machine learning models or techniques, like CNN and even Natural Language Processing, may be used in the future for the dataset we have been working with.

The Common Bag-of-Words model guesses the term based on the context from the input of numerous words. One word or several words may be entered.

Although two semantics can be selected for a single word, the CBOW model averages the input words' context, i.e., two Apple vectors can be expected. Apple as a fruit comes after Apple as a company.

References

1. Al-Mamun A, Akhter S (2018) Social media bullying detection using machine learning on Bangla text. In: 10th international conference on electrical and computer engineering (ICECE) 2018
2. Hussain MG, Al Mahmud T, Akthar W (2018) An approach to detect abusive bangla text. In: International conference on innovation in engineering and technology (ICIET), 27–29 Dec 2018
3. Ahmed FM, Mahmud Z, Biash ZT, Ryen AAN, Hossain A, Faisal Bin Ashraf R (2021) Bangla text dataset and exploratory analysis for online harassment detection
4. Hani J, Nashaat M, Ahmed M, Emad Z, Amer E, Mohammed A (2019) Social media cyberbullying detection using machine learning (IJACSA). Int J Adv Comput Sci Appl 10(5)
5. Pawar R, Raje RR (2019) Multilingual cyberbullying detection system. In: IEEE international conference on electro information technology (EIT)
6. Haidar B, Chamoun M, Yamout F (2016) Cyberbullying detection: a survey on multilingual techniques. In: European modelling symposium (EMS), pp 165–171

Animal Encroachment Detection in Farms: A Survey

V. K. Aravinda, S. Rakesh, Huma Hussain, Ritika Yarlagadda, and Kavya Reddy

Abstract Wildlife such as elephants, monkeys intruding farms that are meant for harvesting is extremely damaging and can result in injury, loss of life of humans and wildlife, damage to human property and crops. Farmers not only lose the harvest that they worked on for months, but it is also dangerous for other livestock and people living in the same vicinity. Farmers resort electrical fencing or barbed wires to avoid this which violate the Prevention of Cruelty to Animals Act. Manually guarding the farm is also not optimal as it takes a lot of effort and there is only limited land a person can cover. This increased the need for an application that will detect animal movement or any unusual activity and set out alarms that can avoid loss of money and effort. Using deep learning and computer vision will help process unstructured data such as images collected at different timestamps. Computer vision using Artificial Intelligence will compare colours and contours in images and indicate an encroachment, giving the owner a much more elaborate recognition of the trespasser. Bearing in mind how most of the farmers are relatively alien to a system of this sort, it requires least amount of intervention from the user as input retrieval and processing is automated. This application could also be incorporated in other areas like to detect wildfires in forests or fields that are often detected late.

Keywords Computer vision · Trespasser · Artificial intelligence · Contours

---

V. K. Aravinda · S. Rakesh · H. Hussain (✉) · R. Yarlagadda · K. Reddy
Department of IT, Chaitanya Bharathi Institute of Technology, Hyderabad, India
e-mail: ugs19067_it.huma@cbit.ac.in

V. K. Aravinda
e-mail: krishnaaravinda_it@cbit.ac.in

S. Rakesh
e-mail: srakesh_it@cbit.ac.in

R. Yarlagadda
e-mail: ritika.yarlagadda@gmail.com; ugs19076_it.ritika@cbit.ac.in

K. Reddy
e-mail: ugs19071_it.kavya@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_44

1 Introduction

Every year, either on highways, roads, sanctuaries, zoos and farms, human–animal interactions have proven to have caused immense losses. These interactions are most probable when wild animals such as elephants and cattle often encroach into weakly protected farms during the night trampling down the crops in the farms. Both animals and humans are put in danger in this scenario and all resources are spoiled. It is not feasible to monitor many acres of land and keep wild animals away from farms. Measures to minimize any probable loss to the farmers, protect crops and the humans on the farm as well must be taken. This can be done using computer vision which is the most used tool with a wide application in video and image processing specially to incorporate interaction involving machines and people. A monitoring system with the least amount of hardware can be implemented through the concept of computer vision. The existing applications in this domain involve hardware and microcontrollers (motors, sensors) which are too exorbitant and inaccessible for the farmers. Thus, exploring a software solution is comparatively less consuming in terms of energy and finance.

2 Literature Survey

"Analysis of video surveillance images using computer vision in a controlled security environment" [1] enlisted methods of recognizing objects comprising facial features in various multimedia input forms such as images and videos. It depicted the use of Haar Cascade from OpenCV to distinguish relevant features from irrelevant ones even during conditions of poor lighting or hardware issues. "Anomalous Motion Detection on Highway Using Deep Learning" [2] presented an anomaly detection dataset which outlined different traffic patterns from vehicles on highways, thus noting down all common and uncommon items. This can be utilized to detect objects that are rather unusual on the roads. "Motion Detection Using Image Processing" [3] proposed a method to spot any movement in the given input using image subtraction and preprocessing. The proposed method uses two subsequent pictures and senses motion in the time interval between the two pictures. It finally eliminates irrelevant features from the two making it useful in surveillance. "Motion Based Animal Detection in Aerial Videos" [4] comprises a video dataset wherein animals perform the most obvious movements in a forest or man-made environment. Optical flow applications are used to evaluate the motion vectors of each pixel of the video. A pixel velocity threshold is used to remove parts of the background and other irrelevant object present. "Body Motion Detection Technology in Video" [5] talks about a few motion detection algorithms from the OpenCV library. Frame Difference method converts each frame captured into grayscale and calculates the median frame. Every frame is then compared to the median to detect a motion. This is combined with edge detection using discontinuity to identify and track motion. "IoT based Anti-poaching and Fire Alarm System for Forest" [6] talks about the connectivity between devices to implement an alert system. LDR sensors can be deployed that detect variation in the amount of light that falls on it. This paper suggests using an Arudino board that sends sensor data to the cloud for regular monitoring. "Efficient Real-Time Object Detection based on Convolutional Neural Network" [7] emphasizes a more user-friendly version of CNN using deep learning and a basic hardware setup based on the Raspberry Pi, along with a compatible camera. Amongst ten predefined classes, it recognizes the distinctive objects within the pictures it takes. "Coarse-to-fine sample-based background subtraction for moving object detection" [8] highlights the drawbacks of conventional background subtraction that cannot function accurately in cases of dynamic background change or illumination. "Enhancing Object Clarity In Single Channel Night Vision Images Using Deep Reinforcement Learning" [9] deals with enhancing the clarity of images that are captured during night. Using feedback from the pixel prediction, q-table is built and images with low resolution are enhanced.

3 Methodology and Approaches

A. *Recognizing trespasser using Adaboost and MongoDB*

As shown in Fig. 1, a record of all the valid users is recorded in a MongoDB database and API face. The video camera records the face of the individual trying to access the installation and compares it to the API face data to look for a match. An alert is delivered to mail using the SMTP protocol if there is no match. Self-learning meta-algorithms that enable the categorization of positive photographs (with a face) and negative images (without a face) using Adaptive Boosting are used to find matches (Adaboost). The Python programming language uses the Haar Cascade algorithms from the OpenCV package to create, delete and transfer information to the user.

B. *Predicting optical flow of the frames*

In this method, generative models are used to recognize a normal motion so that anything that does not match this can be classified as anomalous motion. Conditional generative adversarial networks use labels during the training process to predict optical flow between a sequence of images and consider it as normal movement. The discrepancy between the expected optical flow from the generator and the actual optical flow is used in the second method to identify anomalous motion. A sliding window is then used to average the difference. If there is an inaccuracy in the x or y component of a frame that is greater than a threshold, it is classified as anomalous.

C. *Background subtraction and thresholding*

To be more accommodating of the uncertain environment of the image capture, this method uses homomorphic filtering to enhance illumination in images. It depends solely on the intensity levels and is independent of the hue and saturation values of

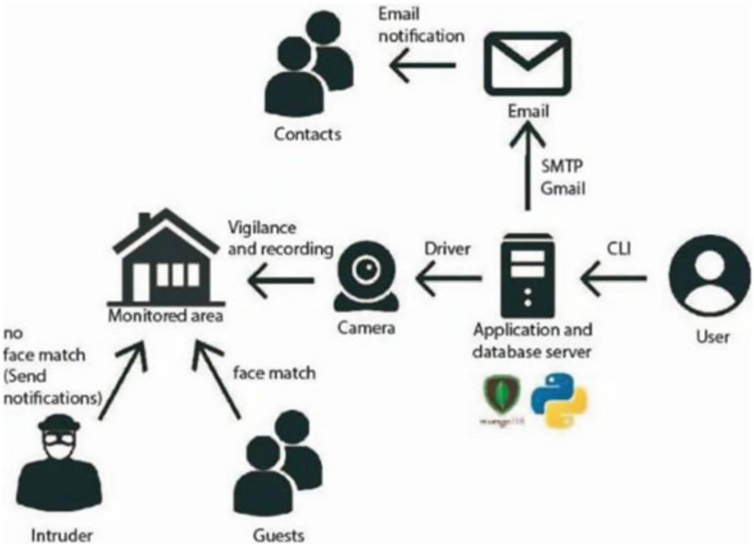


Fig. 1 Facial capture and motion detection flow chart

the objects; the generated images are first transformed to grayscale before further processing. The initial image that is recorded is treated as the background, and each succeeding image is treated as a separate frame. By pixel-by-pixel comparing two frames, it is possible to create a difference image. Thresholding is done on this processed image to obtain insensitivity to noise. To detect an animal, optical flow vectors are used to recognize moving animals based on relative difference between the background and foreground. Motion difference is estimated from the largest velocities of the foreground and background, and the mean value will ignore the unnecessary areas.

D. ***Inter-frame difference and dynamic edge preservation***

This approach is suitable when the camera is static, and the environment shows least changes. A binary image can be obtained by fixing a threshold. Any pixel with the binary value 1 is considered foreground and the rest are background. The relevant objects that need to be detected are in the foreground. For an image sequence that keeps changing, the change in grey value due to motion is detected and a dynamic edge preservation method is implemented. Optical flow vector and optical flow fields are used to detect instantaneous change rate in the grey scale level in a specific area. Each pixel has an optical flow vector to detect motion amongst adjacent frames to finally detect the object. These are shown in Fig. 2.

E. ***Usage of pyroelectric sensors and microcomputer***

This approach is focused on the usage of pyroelectric infrared sensors as the foundation of the design of the system to monitor the daily operation of a laboratory. When

Fig. 2 Edge detection, sceptical descriptors and global features methods

a thief is identified by the infrared detection circuit, the sensor sends the signal to a single-chip microcomputer. This microcomputer then makes an alarm decision. This decision is only taken based on the detection of a human body signal with an infrared probe. The alarm is launched as a buzzer sound equipped with a flashing light acting as an indicator. The main program is accountable for the detection of heat signals of the human body. Through this design, a high-level signal is captured to detect the presence of a human in a laboratory and avoid thefts.

F. ***Aggregating data from sensors using AWS***

The proposed system helps collect information about any intrusion detection and occurrence of forest fire using WSNs. The goal is to outline a remote sensor module that will in hindsight be a piece of the Arduino board. The three modules comprise tree, cloud and authorized officer unit. The tree unit uses a temperature sensor (LM335), smoke sensor (MQ2) and light detective resistor (LDR). The cloud unit uses AWS platform, and the data from Wi-Fi mobile are sent to the cloud via APIs. If the values sensed exceed a fixed threshold, the message is sent through SMS gateway with an alert, device ID and location.

G. ***Feature extraction using CNN***

Transient Search Optimization algorithm can be used to develop variations of feature extraction methods like TSODE. It also uses deep learning to extract features and patterns from raw data rather than depending on supervised machine learning models. CNN model with convolutional blocks is used to identify features. Then a combination of Transient Search optimization and Metaheuristic Search optimization are used to classify the attack types.

H. ***Classification of the detected disturbance***

Pascal VOC and COCO data sets are used for classifying this binary model or to give output as (person, not person). The CNN model used here has feature maps or filters as layers. ReLu layer is also used to alleviate the vanishing gradient problem while processing images. The model with ten classes was trained three times for different kernel filter sizes (3 × 3) and (5 × 5) in convolution layer for 10 and 100 iterations.

I. ***Coarse-to-fine sample-based algorithm***

To solve image clarity issues such as illumination changes, dynamic background, ghosts, concealed and static foreground, the first step is to calculate the region where motion occurs using a combination of images after frame difference is done on them. The second step uses the Vibe algorithm to refine the rough motion region to identify what object has caused the disturbance.

J. ***Optimizing pixel prediction using q-table***

It employed several unique models to boost productivity and adaptability. The first model, called Q-learning, makes use of a Fully Convolutional Network (FCN), a specialized architecture primarily employed in semantic segmentation. The second model employed is A3C, an asynchronous advantage actor-critical algorithm in which many distinct agents each have their own configuration of network parameters. The third model, called PixelRL, is a method of learning in which each pixel has a unique agent, which alters the value of each pixel by engaging in an activity with an effective learning mechanism that significantly improves the presentation by considering both the future conditions of one's own pixel and those of its neighbours.

4 Result Analysis

It was observed that presence of shadows and absence of light resulted in more false positives. Classification using OpenCV fails if the angle of image capture is not correct; hence, motion detection cannot be done without preprocessing the image. The window size of analysis should be varied according to where the object is in the captured image. The pred net model was used to predict up to 6 frames into the future, and it was observed that 5th frame provided better result. Only those images in which significant disturbance is detected must be analysed instead of every frame captured. Duplicate images must be eliminated in case capturing frequency of the camera fails. Scenarios like moving camera, dynamic background and aerial perspective of the scenes were dealt with. Global pixel motion difference between background and animal was detected by highlighting motion features in tracking. The project can further be improved by selecting more appropriate thresholds, having separate colour segmentation for inert animals and combining textures. It was observed that the application of differential evolution algorithms has significantly improved the performance of search algorithms for intrusion. It showcases the usage of deep learning and metaheuristic optimization algorithms to detect anomalous motion. Algorithm built using CNN can recognize up to 10 objects with a rate of up to 98% within a span of 2 s. This was however possible because training set images were preprocessed to cater to unique size and annotation labels. Despite the good response time, its drawbacks include inability to process big data with numerous varieties.

5 Applications

User can monitor large fields, forests, gardens, yards from anywhere. Its implementation can be extended to many avenues including large unmanned and difficult to supervise areas such as—monitoring forests for animal extinction, runways for obstructions, drylands for fires due to heat, railway tracks for protestors/humans. The process of identifying animals in man-made environments can also be used in zoos, wildlife sanctuaries, etc.

6 Conclusion, Challenges and Future Scope

One of the important aspects of the project which is also the issue common to most of these papers is recognizing objects with all the variable features and parameters in place. The techniques discussed in a few of these include object detection, segmentation, tracking and edge detection. TSOE, VIBE algorithm and Haar feature-based cascade classifier are algorithms discussed for the identification and classification of the capture. Connecting sensors like LDR, heat and smoke-sensitive resistors using IoT was analysed, and drawbacks were noted. Challenges to the application pertain to preprocessing the streaming video. Images captured in extremely low illumination and/or distorted images, damaged recording device are few other challenging aspects. In the future, the application can also detect smaller animals and rodents which are camouflaged in crops and harder to detect.

References

1. Casanova G, Yandún D, Guerrero G (2020) Analysis of video surveillance images using computer vision in a controlled security environment. In: 2020 15th Iberian conference on information systems and technologies (CISTI), June 2020
2. Singh H, Hand EM, Alexis K (2020) Anomalous motion detection on highway using deep learning. In: 2020 IEEE international conference on image processing (ICIP), Oct 2020
3. Roy S (2020) Motion detection using image processing. In: 2019 global conference for advancement in technology (GCAT), Feb 2020
4. Fatani A, Abd Elaziz M, Dahou A, Al-Qaness MAA (2021) IoT intrusion detection system using deep learning and enhanced transient search optimization. IEEE Access 9
5. Abd Shehab M, Al-Gizi A, Swadi SM (2021) Efficient real-time object detection based on convolutional neural network. In: 2021 international conference on applied and theoretical electricity (ICATE), May 2021
6. Xu Y, Ji H, Zhang W (2020) Coarse-to-fine sample-based background subtraction for moving object detection
7. Fang Y, Du S, Abdoola R, Djouani K, Richards C (2020) Motion based animal detection in aerial videos

8. Qin L, Liu Z (2021) Body motion detection technology in video. In: 2021 3rd international conference on robotics and computer vision (ICRCV), Aug 2021
9. Ma B (2020) Design of intelligent burglar alarm system in laboratory based on embedded system. In: 2020 13th international congress on image and signal processing, biomedical engineering, and informatics (CISP-BMEI), Oct 2020

The Prediction of Attacks and Challenges in Cyber Security Using Machine Learning Techniques

B. Rekha, S. Rajeshwari, S. R. Siddanna, and G. Kavya

Abstract Cyber-attacks that are automated and persistent are now more likely to succeed in the cyber sphere. The security measures used to identify and respond to cyber-attacks are improved by means of cyber security techniques. Due to hackers' increased intelligence and ability to circumvent traditional security measures, previously utilized protection methods are no longer enough and conventional security solutions are ineffective in spotting new threats and polymorphs. In many cyber security applications, machine learning techniques are essential. Nevertheless, despite the continuous effectiveness of security attacks, there remain considerable difficulties in guaranteeing the reliability of machine learning systems. Main aim of this paper is to offer a literature on machine learning techniques for cyber security in zero-day attack, man in middle, unsolicited mail detection, and malware detection on networks. List of regularly used machine learning tools, concise definitions of each machine learning technique and evaluation standards for classification model evaluation are also included. In addition to providing current information on machine learning trends in cyber security, this study also highlights outstanding problems and challenges in cyber security.

Keywords Machine learning · Cyber security · Cyber-attacks

B. Rekha · S. Rajeshwari (✉) · S. R. Siddanna
Department of Information Science and Engineering, SJB Institute of Technology, Bangalore 560060, India
e-mail: srajeshwari@sjbit.edu.in

B. Rekha
e-mail: brekha@sjbit.edu.in

S. R. Siddanna
e-mail: srsiddanna@sjbit.edu.in

G. Kavya
Department of Computer Science and Engineering, SJB Institute of Technology, Bangalore 560060, India
e-mail: kavyag@sjbit.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_45

1 Introduction

Data transmission over the network from one node to another is the main function of the Internet. Internet usage has significantly expanded as new computer systems, networks, and mobile devices have been created. As a result, cybercriminals and adversaries have turned their attention to the Internet [1]. Machine learning [ML] and artificial intelligence [AI] approaches are being used in variety of fields, including finance [2, 3], education [4], medical [5, 6], manufacturing [7], and especially cyber security [8, 9]. ML approaches are crucial for the early identification and prediction of various threats in many cyber security applications, including spam categorization [10, 11], fraud detection [12, 13], malware detection, phishing, dark web or deep web sites, and intrusion detection. To recognize and counter the cyber-attacks of the new generation, proactive methods are needed.

Cyber-attacks are the most harmful and hazardous, since they have the potential to reveal the most sensitive personal information through phishing technique and also revel the highly classified information of government agencies. This paper gives intensive information of basic cyber security attacks, threats, techniques and tools for ML, current scenario of ML in cyber security, and challenges using ML tools for cyber security. This paper considerably differs from the sooner articles in those respects.

The contributions of this survey article are:

- A discussion of the primary cyber security attacks and accompanying responses.
- Complete details on widely used ML tools and techniques.
- A discussion on current state of ML in cyber security.
- A discussion on techniques for cyber security by using ML.

2 Types of Cyber-Attacks

The fundamental concepts relating to cyber-attacks are introduced in this section. A cyber-attack now without a doubt comprises disrupting computer operations on the victim's network or gaining access to the victim's computer without authorization. Depending at the perspective, there are various classes into which cyber assaults may be subdivided [13], which are given below.

2.1 Zero-Day Attack

Zero-day attacks are cyber-attacks when the attacker exploits system weaknesses to acquire unrestricted access to the system in order to insert malware or spyware [14]. This suggests that there is no widely accessible patch and it also implies that the vendor may or may not be aware of it. As suggested by the names, zero-day

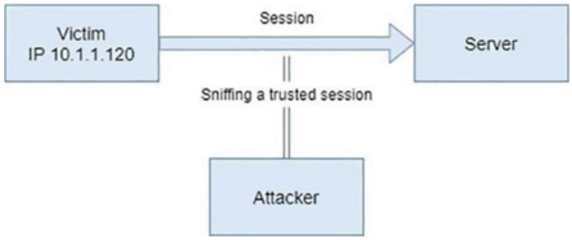


Fig. 1 Instance of a man-in-the-middle attack [14]

attack and zero-day exploit, the software developer has "zero-day" to fix the system's vulnerability. An antivirus program cannot identify a new zero-day assault using a signature-based technique since it launches its attack on the day it is formed due to a system flaw [14].

2.2 Man-in-the-Middle Attack (MitM)

This assault takes place when a hacker bridges a secure connection between a client and server. Session hijacking is a best example of a MitM attack. Here, a victim and the server are engaged in a session when an attacker hijacks the session and attacker replaces the victim's Internet Protocol (IP) with their preferred IP and keeps the session going with the server in this situation, the server views the attacker's IP.

During this process, the attacker's computer spoofs the victim's organization information and sequence number as a reliable client [14]. While the victim's PC is disconnected as shown in Fig. 1.

2.3 Web Access Compromise

This assault takes place when vulnerabilities' exploiting on websites. The structured query language (SQL) injection attack and cross-site scripting (XSS) are two common types of web-compromising attacks [2].

SQL Injection Attack: Here, attackers enter data from the client to the server, such as login credentials, and then write SQL queries into the database, which affects database-driven websites. So, the attackers utilize a preset SQL statement rather than the anticipated data for a post request. This command operates, reads touchy statistics from the database, from time to time has the capacity to edit touchy statistics, and plays administrative obligations, if the database has read most effective access.

XSS Attack: Attackers use third-party websites to force users to run or download scripts into their web browsers. JavaScript code with a payload is frequently injected by attackers into several websites. The victim's Internet browser processes

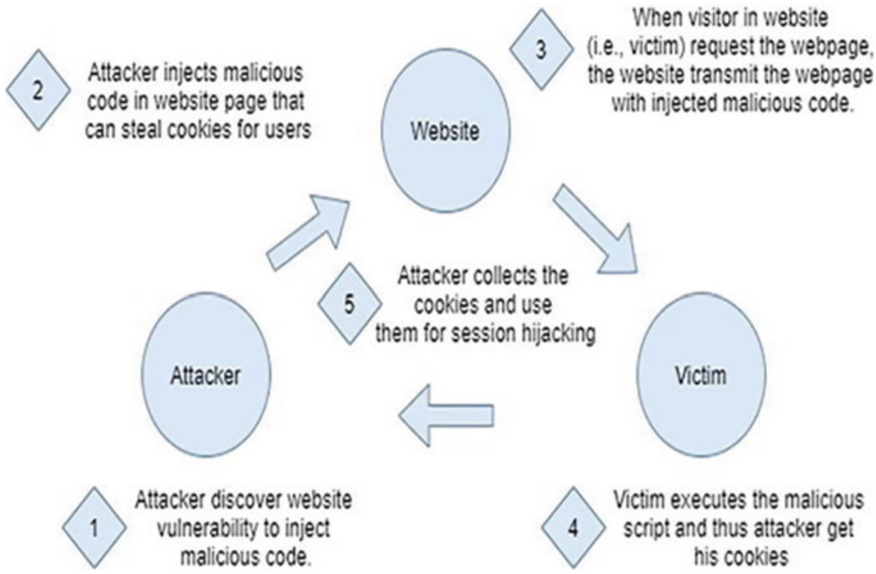


Fig. 2 Cross-site scripting attack [2]

the attacker's script and payload when the customer searches or requests information from the websites.

The victim's session cookie can be unlocked by the attacker's malicious script, which can then be used to steal information like keystroke logs. Additionally, it permits the attacker to get admission to the victim's computer. An overview of XSS is given in Fig. 2.

2.4 Malware Attack

Unwanted software is referred to as malware (short for malicious software). Cyber criminals have long utilized malware to accomplish their objectives, including deleting or turning off a cyber-physical system, collecting private, extensive data, infiltrating a network or system, and inserting dangerous scripts [13]. Malware may be damaged down into some of exceptional bureaucracy relying at the targets of the attackers and the way regularly they spread. They frequently include viruses, worms, Trojan horses, spyware, ransomware, scare ware, bots, and root kits.

Viruses: Viruses are hostile pieces of code that travel with host software and corrupt files on both the host computer and the shared network, much like a living virus in the human body. The Melissa virus and the Creeper virus are two best examples of harmful viruses [11].

Worms: Worms have a different method of spread than viruses. Worms can spread without a host machine, unlike viruses. Email attachments frequently include self-replicating worms [10]. Additionally, worms do not harm host computer files. By copying themselves to every contact on the victim's electronic mail account and utilizing the available network resources, worms can launch a denial-of-company attack. The common examples of worms are SQL Slammer and Code Red.

Trojan: A Trojan differs greatly from viruses and worms in terms of what it is intended to do. To fool people into installing a Trojan on their computer, attackers employ social engineering techniques. Unlike viruses and worms, Trojans do not replicate themselves or infect host computer files; instead they build a backdoor that allows attackers to run dangerous software when necessary. Zeus and She dun Android Viruses are some examples of Trojans [10].

Spyware: Instead of instantly initiating an assault, spyware is employed to monitor user activity. Without the consumer's understanding or agreement, this program is used to infuse borrow touchy consumer data, such as login credentials and keystroke records [9].

Ransomware: This sort of malware differs from others in that its payload establishes a procedure to call for a ransom from the sufferer, similarly to infecting the sufferer's files. Typically, Trojans are used to carry out ransomware assaults. The common example of ransomware includes like WannaCry, Torrent Locker, and so on [12].

2.5 Denial of Service (DoS)

The crucial cause of this sort of cyber-attack is to absolutely smash a system's or network's ordinary operational statistics. There are three primary types of DoS attacks.

Based on Host: In host-based attacks, malicious software program or personal computer worms are established on host machines after which used to perform their payload or operation to flood the whole community—beginning with the host—with account less flow of host requests.

Based on Network: Attacker chooses to run their payload on an entire network as opposed to a host machine in order to disrupt the network's regular operations.

Distributed: A Distributed Denial of Service (DDoS) attack is typically launched from more than a few tools and from a network if you want to absolutely near down the victim's network.

3 Techniques and Tools for Machine Learning

Machine Learning Basics

AI is a subfield of computer science that creates methods, theories, and software. Cyber protection and early detection of some of novel and automatic attacks [5, 6], in addition to the identity of phishing websites is given. Supervised gadget mastering, unsupervised gadget mastering, and semi-supervised gadget mastering are the three most important agencies that gadget mastering can be divided into primarily based totally on technique. In supervised gadget mastering, the intended labels or modules are already known for the data, e.g., classification and regression. The targeted value is unknown in unsupervised machine learning. Unsupervised studying is mostly worried with figuring out styles with inside the records. It features via way of means of figuring out styles within side the records, including clustering. Semi-supervised gadget studying is used, while several records are categorized or while human experts are required all through the records series process. Unquestionably, the human professional all through the labeling way will assist to deal with the trouble and enhance the model's accuracy [4].

Reinforcement mastering (RM) is a unique subfield of gadget mastering. Due to the fact that the algorithms achieve remarks for any incorrect predictions, RM is also known as learning with a critic. However, the set of rules has not been advised of a way to restoration [7]. Based on a gadget of rewards and penalties, this phenomenon operates. Alpha Go is a well-known illustration of this method [3]. The field of cyber security uses deep reinforcement learning. A part of system mastering is known as deep learning (DL). Though each deep mastering and gadget mastering rent the equal strategies and tasks, they have got distinct capacities. The two primary DL research fields are deep belief networks and convolution neural networks. Over the past ten years, these areas have drawn the attention of the academic and research community. Today, an example of DL is automatic driving. Numerous studies that used models to improve cyber security have been reported in the literature.

4 Commonly Used Machine Learning Techniques

The machine learning approaches which are commonly used are described in this section. A brief summary of ML models, including their temporal complexity, is described below [1].

4.1 Support Vector Machine

The support vector machine (SVM) is notion to be the maximum well-appreciated and green system learning (ML) technique. SVM divides and categorizes the two data classes. The margin and separation between the hyper planes can be raised to improve classification accuracy. The records factors alongside the threshold of the hyper aircraft are referred to as aid vector factors. SVM is divided into top groups. The kernel function determines whether it is linear or nonlinear. The processing reminiscence and schooling time necessities for SVM are high. SVM calls for schooling at more than one time periods to examine the dynamic user's conduct for higher results.

4.2 Decision Tree

Supervised machine learning methods based on recursive tree architectures include decision trees (DT). DT is made up of three parts: A thing or quality is represented by a tree's root or intermediate nodes. The capacity values of the determine node are represented with the aid of using every divergence route within side the tree (object). An instance of the projected category/labeled feature is the leaf node. If then regulations are some other techniques for representing the ensuing tree, the optimum additional intermediate node is chosen during tree construction using assessments of entropy and information gain. CART [1], decision tree algorithms that are thought to be noteworthy are C4.5 [14] and ID3 [2]. ID3 operates via a greedy mechanism. But it is unable to manage attributes with numerical values. ID3 is an improved version that addresses the problem of over fitting by using techniques for tree with noisy data as an exception, it can handle the over fitting issue. Categorical and numeric attributes are supported by CART, and it also accepts missing data that ID3 cannot handle.

4.3 **K-***Nearest Neighbor*

A way for unsupervised studying is *K*-nearest neighbor (kNN). At its foundation, the gap characteristic establishes how numerous or comparable records times are. Compared to different classifiers, this one requires much less schooling time. However, due to the computing time involved, the categorization process takes of kNN may be differentiated.

The two different ways of computing the oddity scores depend on the (1) distinction between the kth neighbor and the piece of information, not entirely set in stone. (2) The thickness of every information occasion is considered in the computation. The *k*th information point's worth affects the classifier's general presentation. The decision of distance capability to decide the partition/distinction between relevant pieces of information and loud information both influence how well the classifier performs. KNN is exorbitant to figure and needs a great deal of capacity to be controlled. The distance between information focuses x and y is regularly resolved utilizing the Euclidean distance equation, $d(x, y)$.

4.4 Random Forest

Random forest (RF) is an illustration of group realizing, which consolidates different classifiers to make an issue speculation and a run of the mill result. It is utilized for grouping and relapse applications and is otherwise called an irregular choice timber land. RF is seen as a superior emphasis of truck. RF is in many cases an assortment of estimated results from different choice trees. In the writing, the irregular woodland has been utilized for interruption location and estimating spam volume [9, 10]. It performs better on non-straight issues and uses less computational assets during the model's preparation stage.

4.5 Naive Bayes

The Naive Bayes (NB) class of classifiers uses the Bayes' theorem (also known as the Bayes' Rule) to deconstruct the conditional probability of a problem under consideration. Different attack types do not, however, meet this independence need in terms of cyber security. A way for unsupervised studying is *K*-nearest neighbor (kNN). At its foundation, the gap characteristic establishes how numerous or comparable records times are. Compared to different classifiers, this one requires much less schooling time. This classifier is stated to be greater sincere and to discover items greater quickly.

Naive Bayes consists of the multinomial, Bernoulli, and Gaussian approaches [8]. The underlying premise of this classifier is that comparable facts factors in an area clustered collectively as opposed to apart. Based on anomaly scores, essential subcategories are treated with the aid of using multinomial Naive Bayes. The range of instances that this occasion happens is represented with the aid of using function vectors in those values [9]. Binary function vector category is performed the usage of Bernoulli Naive Bayes.

A right instance of this method is Bags of Words [11]. Continuous data values are subjected to the classification algorithm Gaussian Naive Bayes. These values are spread out according to the Gaussian distribution [13].

4.6 Artificial Neural Network

ANNs are a chain of ahead by skip and lower back propagation cycles. As information is dispatched ahead, it is miles acquired with the beneficial useful resource of the use of every node of a hidden layer. The activation rate of every node with inside the hidden layer and output layer is determined. The activation characteristic of a classifier impacts how properly it performs. The error is derived by subtracting the goal value from the network output. The Guardian Descent mechanism in back propagation is used to communicate this difference back to the input layer, where it is used to modify the weights between hidden and output nodes. Until the required threshold is reached, this process is repeated.

Although it takes some time to train, ANN is a nonlinear version that is straightforward to apply and is reputed to be noise resistant [10]. Taveras made an effort to analyze the significance of end-user password entry habits for account security.

To reduce the possibility of account hacking, they have recommended changes to password entry practices. They asked the participants to put down any password of their choosing as part of their study. This examination hired system studying strategies to get the predictions, mainly neural networks. Overall, the examiner observed that while neural networks can be efficaciously used to make predictions, there had been nevertheless sure limitations. The fact that users' conduct did not follow a logical development despite the fact that the majority of participants had backgrounds in information technology was a negative [11].

5 Current Scenario of Machine Learning in Cyber Security

According to some cyber security offers, protection from online threats and assaults. The identity and category of malicious URLs, economic fraud, unsolicited mail category, IDS, the development of malicious domains, probing, cyber-extortion, and malware are only some of the various dimensions of cyber security [9]. In addition to targeting computer networks, cyber criminals are increasingly concentrating their attacks on mobile networks and gadgets. As a way we are aware, no survey has ever particularly investigated any issue of an assault that concurrently goals laptop networks and cellular devices.

Cyber security includes characteristics of cyberspace such as network security, ICT security, and Internet security. Three critical regions of cyber security—the detection and category of IDS, spam, and malware—wherein ML strategies are critical have stuck our attention. These risks to laptop networks and cell gadgets were included in extra detail. Intrusion detection systems for computer networks can also be divided into hybrid, anomaly, and signature-based/misuse-based techniques. Subtypes of infiltration can be used on either a host or a computer network, according to further classification [9]. The definition of spam detection is expanded in relation to the media, encompassing images, emails, SMS, videos, and Twitter. Malware is also investigated in terms of static and dynamic analysis. In the literature, ML approaches are used to combat several kinds of cyber-attacks. One potential solution to quickly respond to cyber-attacks is machine learning (ML). Because they are quick to respond to new threats and can learn from prior errors, such issues are handled using ML approaches [10].

6 Challenges of Using Machine Learning Techniques for Cyber Security

Common Problems with Machine Learning Models

The area of cyber safety typically makes use of gadget getting to know strategies. But there are a whole lot of limitations within side the way. ML strategies want an extensive quantity of high-overall performance assets and facts while developing the models. Using numerous GPUs is one option, but this is neither a cost-effective nor a power-efficient solution [3]. Additionally, ML methods are not intended to identify cybercrimes. Traditional ML algorithms did not place much emphasis on cyber security. It is necessary to have strong and reliable ML methods that are especially made to counter security threats and deal with adversarial inputs. It should be noted that different security assaults cannot be effectively detected by a single ML model. To deal with a particular kind of data, a specific ML model needs be developed. Another difficult task is the early attack prevention. These real time and zero-day threats should be able to be quickly detected using ML approaches [5]. Machine getting to know fashions were applied with inside the area of drugs to diagnose ailments and make selections concerning the detection of terrorism. Under these conditions, prediction cannot be blindly depended upon to prevent unfavorable events. Machine gaining knowledge of techniques has to provide a few high-stage correctness assurances in location of velocity and accuracy while applied in life-essential or mission-essential applications (inclusive of self-sufficient vehicles, cyber security, and surgical robotics) [6]. Trusting a prediction, or whether consumer will rely upon a given prediction to take a positive action, and trusting a version, or if a consumer will well rely upon a version used as a tool, are particular methods to signify a classifier‘s dependability.

It is beneficial to understand what wishes to be finished to show a mistaken version right into a dependable one. Traditional linear or shallow mastering is regularly greater dependable, while being slower or much less accurate. Despite a theory that is continually changing, deep learning is still rather difficult and opaque. Cell telephones and the worldwide positioning machine have made it viable for forensic technological know-how and sickness manipulate to discover transferring objects. Maintaining the object's credibility is a hard assignment because of the probability of in accurate or up to date facts on cellular devices. Chenyun published a method to assess the degree of similarity between information about an object's position obtained from diverse sources when establishing how accurate location data produced from moving item trajectories is, there is always a potential of ambiguity. This uncertainty exists because of network latency as well as the fact that the objects are shifting their locations [15].

7 Other Challenges of Using Machine Learning for Cyber Security

The maximum current system studying algorithms and strategies which have been carried out to fight cybercrimes like junk mail and IDS. The exposure of numerous such inconsistencies and problems provides a solid foundation for analyzing additional upcoming difficulties and trends. Below are some of these topics discussed.

7.1 Datasets

For each dataset, there are varying numbers of features and categories. Most attack-related information and content is unnecessary. Machine studying fashions carry out higher, while there is a huge quantity of facts to be had for education, which isn't the case for the datasets which might be presently to have had Benchmarks and fashionable datasets with a huge quantity of facts and a lightly allotted wide variety of assault classes need to be furnished for education and checking out reasons [9].

7.2 Evaluations Metrics

When using the same dataset, the majority of researchers have ignored one aspect of the situation by using alternative parameters to assess a classification model. To make future improvements to the model, it is necessary to take into account a set of agreed upon standard metrics for comparison [9].

7.3 Detection and Time Complexity of Various Techniques

The literature gave little thought to the attack's immediate context. It is also important to take into account an algorithm's temporal complexity and the attack detection rate in a real-time context. Every day, cyber criminals develop new assaults to reveal the network's weaknesses. One very important factor to think about is how effectively an assault can be detected. Security analysts will spend time searching into the coolest behavior. Security professionals will begin to lose religion within side the gadget if it maintains to offer misguided alerts.

8 Conclusion

In order to improve security measures to recognize and respond to cyber-attacks, cyber security has grown to be a concern on a global scale. The security solutions that were previously in use are no longer sufficient since they are incapable of detecting brand-new attacks that take on multiple forms. In many different applications of cyber security systems, machine learning techniques are essential. This paper briefly presents the uses of machine learning techniques in the area of cyber security are briefly discussed. Even the most cutting-edge ML model finds it challenging to deal with these cyber-attacks due to the unique characteristics of each cyber threat. It is impossible to make a single suggestion based on a single model for all attacks. Several aspects, including detection rate, time complexity, classification time to detect new and zero-day attacks, and accuracy of an ML model, should be considered when selecting a specific model to identify a cyber-attack. Additionally, a thorough analysis of cyber security datasets is provided. Finally, a list of common problems and difficulties in applying ML models to cyber security has been provided.

References

1. Bhattacharyya DK, Kalita JK (2013) Network anomaly detection: a machine learning perspective. Chapman & Hall, London
2. Gu S, Kelly BT, Xiu D (2019) Empirical asset pricing via machine learning. In: Proceedings of 31st Australas. Finance Banking Conference, Chicago Booth Res. Paper 18-04, Yale ICF Working Paper 2018-09, Sept 2019 (Online). Available https://ssrn.com/abstract=3159577
3. Emerson S, Kennedy R, O'Shea L, O'Brien J (2019) Trends and applications of machine learning in quantitative finance. In Proceedings of 8th international conference on economics and finance research (ICEFR), 2019, pp 1–9
4. Prasad R, Rohokale V (2020) Artificial intelligence and machine learning in cyber security. In: Cyber security: the lifeline of information and communication technology. Springer, Cham, pp 231–247
5. Apruzzese G, Colajanni M, Ferretti L, Guido A, Marchetti M (2018) On the effectiveness of machine and deep learning for cyber security. In: Proceedings of 10th international conference cyber conflict (CyCon), May 2018, pp 371–390
6. Torres JM, Comesaña CI, GarcíaNieto PJ (2022) Machine learning techniques applied to cyber-security. Int J Mach Learn Cybern 10(10): 2823–2836, 2019 ([29] Difference between threat and attack. Accessed: June 3, 2022)
7. Li B-H, Hou B-C, Yu W-T, Lu X-B, Yang C-W (2017) Applications of artificial intelligence in intelligent manufacturing: a review. Front Inf Technol Electron Eng 18(1):86–96
8. Virmani C, Choudhary T, Pillai A, Rani M (2020) Applications of machine learning in cyber security. In: Handbook of research on machine and deep learning applications for cyber security. IGI Global, Hershey, pp 83–103
9. Shaukat K, Luo S, Varadharajan V, Hameed IA, Chen S, Liu D, Li J (2020) Performance comparison and current challenges of using machine learning techniques in cyber-security. Energies 13(10):2509
10. Almukaynizi M, Grimm A, Nunes E, Shakarian J, Shakarian P (2017) Predicting cyber threats through hacker social networks in darkweb and deepweb forums. In: Proceedings of international conference computional social science Americas (CSS), 2017, pp 1–7

11. Jain AK, Goel D, Agarwal S, Singh Y, Bajaj G (2020) Predicting spam messages using back propagation neural network. Wireless Pers Commun 110(1):403–422
12. Alauthman M, Almomani A, Alweshah M, Omoushd W, Alieyane K (2019) Machine learning for phishing detection and mitigation. In: Machine learning for computer and cyber security: principle, algorithms, and practices. CRC Press, New York, p 26
13. Mubalaike AM, Adali E (2018) Deep learning approach for intelligent financial fraud detection system. In: Proceedings of 3rd international conference on computer science engineering (UBMK), Sept 2018, pp 598–603
14. Thomas T, Vijayaraghavan AP, Emmanuel S (2020) Machine learning and cyber-security. In Machine learning approaches in cyber security analytics. Springer, Singapore, pp 37–47. Available https://link.springer.com/chapter/https://doi.org/10.1007/978-981-15-1706-8_3
15. ICT Fact and Figures 2017. Accessed June 1, 2020 (Online). Available https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf

Stock Forecasting Using Multichannel CNN and Firefly Algorithm

Nilesh B. Korade and Mohd. Zuber

Abstract Investing in stocks is a smart strategy to use your money and ultimately grow your wealth. Your money may grow in value and outpace inflation if you make intelligent investment decisions. Neural networks are frequently used in automated forecasting strategies to forecast stock movement and price. In order to anticipate future stock values, researchers are using convolutional neural networks (CNNs) to extract patterns, also known as features, from time-series data. Compared to other approaches, CNN offers a more promising result and has the potential to anticipate stock prices and movements accurately. Better results are obtained when the training parameters for CNN on a stock dataset are set appropriately. The performance of CNN on the same dataset differs as neural networks are stochastic in nature and use randomness during learning. In this paper, a multichannel convolutional neural network (MCNN) is applied to enhance CNN forecasting accuracy, and a firefly algorithm (FA) is used to find the optimal configurations for CNN architecture. When compared to the state of the art, the results reveal that the proposed method requires fewer iterations to find the optimal parameter and provides improved accuracy.

Keywords Stock · Optimization · Firefly algorithm · Convolutional neural networks · Forecasting

1 Introduction

Although purchasing Indian stocks may seem lucrative, not all investors have the opportunity to receive the most return on their investment. Thousands of Indian stocks will be available to investors, but making the right choice matters [1]. Unfortunately, there isn't a simple formula that can accurately predict how a stock's price will

N. B. Korade (✉) · Mohd. Zuber
Department of Computer Science and Engineering, Madhyanchal, Professional University, Ratibad, Bhopal, Madhya Pradesh 462044, India
e-mail: nilesh.korade.ml@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_46

move. There are many different factors that influence stock prices. Fundamental variables determine stock values based on a company's earnings and profitability from producing and selling goods or services. Technical factors include momentum, chart patterns, investor activity, etc., based on a stock's price history in the market. Because of the variability, predicting the exact stock value is difficult [2]. In particular, deep learning algorithms have been effective at anticipating market movements. A deep learning method known as CNN yields satisfactory results, identifies changing stock price trends, and has a high degree of forecast accuracy [3]. It is necessary to choose certain parameters for training CNN on stock datasets, including epoch, batch size, filters, activation, etc., which affect CNN performance. Due to the stochastic nature of learning, each time CNN is trained on the same dataset, the accuracy varies. Optimizers are algorithms or techniques that adjust the attributes of your neural network to reduce losses without impacting training efficiency [4]. Recent studies have implemented metaheuristic optimization algorithms to address optimization problems by iteratively identifying the best CNN architectural parameters that boost CNN performance. Currently, the best techniques for CNN hyperparameter optimization include the firefly algorithm (FA) [5], random search (RS) [6], particle swarm optimization (PSO) [7], ant colony optimization (ASO) [8], etc. The objective function is a test function that accepts a fixed number of input variables and, upon calculation, yields the input cost. The optimization process seeks the best values for an objective function that reduces or increases cost. The proposed study aims at establishing a hybrid approach, in which the multichannel convolutional neural network (MCCNN) [9] is employed to enhance CNN performance by executing multiple CNNs parallelly in stock forecasting and the firefly algorithm (FA) is employed to identify the optimal configurations for CNN architectures. Each CNN output in the MCNN is averaged to get the final output. The presented method is compared to state-of-the-art methods, and the findings demonstrate that FA picks the right parameter in fewer iterations and MCNN increases predicting performance by reducing error.

2 Literature Survey

Machine learning (ML) techniques learn from past data by identifying stock patterns and stock movements over time, which helps to improve prediction accuracy. A ML model can be trained to predict stock price fluctuations and understand trends based on massive amounts of data. Support vector machines, Naive Bayes, K-nearest neighbor, and random forest are examples of traditional ML techniques. Researchers use data from Yahoo or NSE-India to train ML models and then evaluate the accuracy of each model to choose the best one [10]. With the advancement of deep learning, several neural network-based techniques have been developed that offer promising results for forecasting issues [11]. When a large dataset is provided, the LSTM is a special type of recurrent neural network (RNN) that can recall past input in its memory and forecast the future [12, 13]. CNNs have gained a lot of attention in fields like computer vision, image processing, and others. Recently, there has been an increase in interest among researchers in using CNNs to forecast time-series issues. Since CNN outperforms other deep learning techniques in terms of accuracy at predicting the direction of stock market movements, it can be successfully employed in trading systems [14]. Although CNN has a good reputation for addressing any kind of learning problem, its parameters can make it difficult to train. To overcome the optimization issue in CNN, various optimization techniques were developed to achieve a balance between accuracy degradation and processing speed. The random search (RS) randomly picks a set of variables to train neural networks and assesses their effectiveness against past versions. An enhanced variant of RS called weighted random search, which is intended for hyperparameter optimization, assigns probabilities of change pi, $i = 1 \ldots d$ to each dimension [15]. PSO is a population-based stochastic evolutionary computation technique of the "swarm intelligence" family that was developed by Eberhart and Kennedy and inspired by the social behavior of flocking birds in search of food sources. Individuals are referred to as particles in PSO, and each particle has a specific velocity that represents the moving parameters and position that stand for the attributes of a solution. According to its own and other members' learning experiences, each particle of the population continuously adjusts its search pattern to find the most appropriate solution [16, 17]. A probabilistic method for resolving computational issues that can be reduced to finding efficient routes via graphs is the ant colony optimization algorithm (ACO). To help more simulated "ants" find better solutions during following simulation iterations, the simulated "ants" automatically record their positions and the class of their solutions [18]. The firefly algorithm is a population-based metaheuristic optimization algorithm used in mathematics that is inspired by the flashing behavior of fireflies and simulates the firefly's attraction to flashing light. Each firefly has a position, and the less dazzling ones will move toward the brighter ones by altering their position [19]. By combining multiple kinds of essential neural networks such as an ANN, CNN, or RNN, splitting different computational flows, and integrating them into one output, multichannel input enables us to construct a single neural network. The use of multichannel CNN enables the merging of various findings from individual CNNs into a single output that is more accurate. Comparing the MCNN to other neural network approaches, significant performance improvements are seen [20, 21].

3 Methodology

3.1 Dataset Description

Pandas Datareader, a Python module, was used to gather information on Infosys stock from Yahoo Finance between January 1, 2003 and October 31, 2022. The Infosys Stock dataset attributes and their values are mentioned in Table 1. We picked the closing price of a stock as the target variable since it is possible to estimate the profit or loss using the stock closing price on that particular day.

Table 1 Infosys stock dataset

Date	High	Low	Open	Close	Volume	Adj. close
1/1/2003	75.39	74.61	74.69	74.99	32,066,112	51.65
2/1/2003	75.60	73.59	75.53	73.79	42,260,032	50.83
3/1/2003	75.38	74.22	74.84	75.00	60,483,072	51.67
…	…	…	…	…	…	…
27/10/2022	1534.90	1512.50	1530.80	1523.95	6,760,137	1523.95
28/10/2022	1527.95	1507.00	1526.00	1513.25	3,053,981	1513.25
31/10/2022	1546.40	1530.25	1530.25	1537.65	4,542,336	1537.65

3.2 CNN and MCNN

CNNs have emerged as the go-to technique for resolving every image data issue, and their use in video analytics is also growing. In particular for text classification, natural language processing (NLP), or forecasting problems, CNN gives outstanding performance. In order to effectively apply CNN for stock forecasting, one-dimensional time-series data must be transformed into an input image matrix structure. The fundamental layers of a 1D-CNN are the convolution layer (Conv1D), max pooling layer (MaxPooling1D), flattening layer (Flatten), and dense layer that includes a fully connected and output layer, as illustrated in Fig. 1. Convolution operations are carried out by the convolution layer by accepting input data and applying filters over it. The final output is the dot product of the filter and the input matrix. When performing convolution, padding stretches the processing region by adding pixels to the input matrix, and stride controls how the filter is moved over the input matrix. Padding and stride operations can be used to adjust the height and width of result vectors. In order to increase computational capacity by shrinking the dimension of the data, the pooling layer multiplies the resulting matrix from the convolution layer and the pooling matrix. The dropout layer, which removes a few neurons from the NN during the training phase, can remove overfitting, a typical issue in neural networks. The final few layers that define the output are fully connected layers that receive one-dimensional series created by the flattening operation as input. The activation function is the nonlinear transformation we apply to the input signal to make it more accurately fit the outcomes. The most common activation functions in NN are the relu, logistic or sigmoid, tanh, etc.

By integrating several forms of fundamental neural networks, such as an ANN, CNN, or RNN, dividing up distinct computational flows, and combining them into a single output, the multichannel technique makes it possible to create a single neural network. In contrast to individual model predictions, which may be incorrect due to the stochastic nature of neural networks, MCNN performs an average of values predicted by various underlying CNN models. By taking into account, the results from various neural networks, predicting accuracy can be improved. Each channel in MCNN consists of the convolution, max pooling, flattening, dense layers, etc., that are shown in Fig. 2.

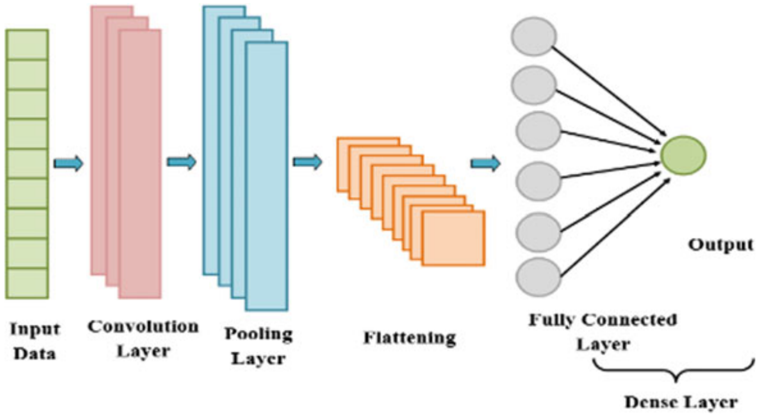


Fig. 1 CNN architecture

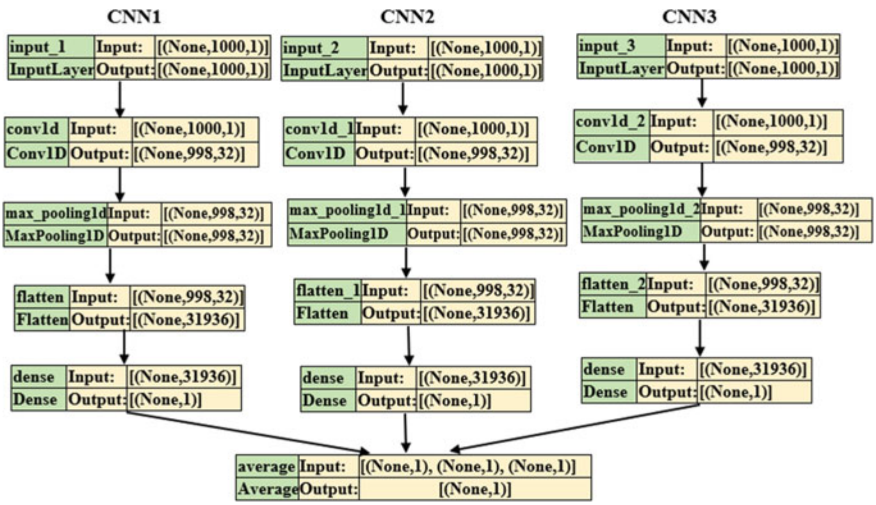


Fig. 2 Multichannel CNN

3.3 Firefly Algorithm

The firefly optimization algorithm is Xin She Yang's 2008 invention, inspired by how fireflies' mate. There are three factors that make up an optimization problem: the function to be optimized, the optimization rule, and the collection of possible solutions from which to choose the value for the variable. The brilliance of the firefly is determined by the objective function, whose value is to be either reduced or enhanced. By altering their position, a firefly with lower brightness will move toward a firefly with higher brightness, regardless of their sex, as all fireflies are unisex. Their brightness determines how attractive they are, and attractiveness reduces as distance grows between them.

Inputs to FA: 'α'—randomization parameter. 'β_0'—attraction, 'Υ'—absorption coefficient, '$f(X)$'—objective function, variable boundary, 'n'—population size, 'd'—number of dimensions, 'it'—number of iterations.
Output: the best parameter.
Step 1. Use the following equation to determine each firefly's random position X in all dimensions.

$$\text{Position }(X_i) = \text{lower\_bound} + \text{random\_number} \\ * (\text{upper\_bound} - \text{lower\_bound}) \quad (1)$$

Step 2. Find the cost or brilliance of each firefly using the objective function $f(X_i)$, where $i = 1$ to n.
Step 3. For it $= 1$ to iter.
Step 4. For $\text{ff}_i = 1$ to n.
Step 5. For $\text{ff}_j = 1$ to n.
Step 6. If $f(X\text{ff}_i) > f(X\text{ff}_j)$.
Step 7. Move the firefly ff_i in the direction of the firefly ff_j after calculating the new position and cost.
Step 8. If not, firefly ff_i will move randomly.
Step 9. End if.
Step 10. End for ff_j.
Step 11. End for ff_i.
Step 12. Choose the most cost-effective position.
Step 13. End for it.

The brightness of the firefly is determined by the objective function; if the MSE is lower, the firefly will be brighter, and so on. Using Eq. 3, the less bright firefly i will move in the direction of the more bright firefly j.

$$X_i^{it+1} = X_i^{it} + \beta_0 e^{-\Upsilon r_{ij}^2} * \left(X_j^{it} - X_i^{it}\right) + \alpha_{it}\,\epsilon_i^{it} \quad (2)$$

The ith firefly's present position is indicated by the X_i^{it} and the updated position by the X_i^{it+1}. A Gaussian distribution was used to obtain the random number ϵ_i and the range of the light absorption coefficient (γ) is 0.01 to 100. The randomization parameter is α_{it}, and the attraction parameter is β_0, with $\beta_0 = 1$ and $\alpha_{it} = 0$ being the recommended values. The distance between two fireflies, i and j, is calculated using Euclidean distance as shown in Eq. 3.

$$\text{Euclidean distance }\left(r_{i,j}\right) = \sqrt{\sum_{z=1}^{d}\left(X_{i,z} - X_{j,z}\right)^2} \quad (3)$$

Figure 3 illustrates the proposed FA-MCNN flowchart. MinMaxScalar was applied to convert values into a similar range on the dataset downloaded from Yahoo Finance. The FA determines the best parameter for CNN, and the resulting value is then used for training MCNN to calculate the accuracy of the proposed method during training and testing.

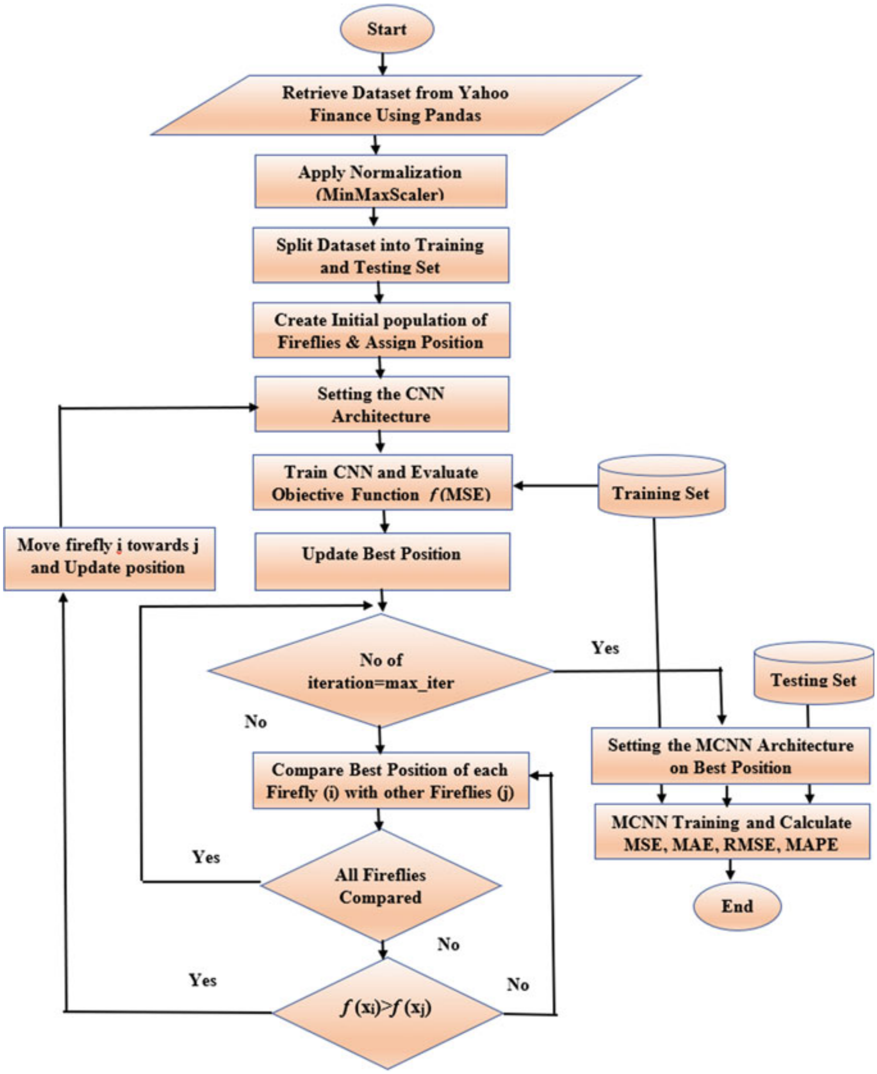


Fig. 3 Flowchart of proposed methodology

4 Results

Infosys stock dataset divided into training set contain 4904 records from January 1, 2003 to September 30, 2022 and testing set contain 19 records from October 1, 2022 to October 31, 2022. The aim of the objective function is to reduce MSE. The objective function, which receives a parameter from the optimization algorithm, trains CNN and returns the computed MSE. The proposed FA-MCNN's performance was compared to that of the RS-CNN, PSO-CNN, and FA-CNN. Different evaluation metrics such as MSE, MAE, RMSE, and MAPE are used to access the performance of the CNN model trained on different optimization techniques. Finding the best option from a range of potential alternatives is the aim of an optimization issue. The outcome may differ if we use CNN architectures with different parameter values to solve the same problem; hence, choosing the proper hyperparameter improves accuracy with a shorter epoch. The hyperparameter boundary values and a list of the control parameters for RS-CNN, PSO-CNN, FA-CNN, and FA-MCNN are given in Tables 2 and 3.

The input shape (1000, 1) is used to train the model, which uses the previous 1000 records to predict the stock price for the next day. While calculating the October month's stock price, previous 1000 records were used to calculate the next day's price, then the forecasted price and previous records, and so on. The training and testing performance for stock price forecasting using CNN trained on parameter return by various optimization techniques is discussed in Table 4.

Table 2 Upper and lower bounds for the hyperparameters

Parameter	Filters	Kernel size	Pool size	Batch size
Lower bound	16	1	1	32
Upper bound	128	5	5	256
Step size	16	1	1	32

Table 3 List of the control parameters

Control parameter	PSO-CNN	Control parameter	FA-CNN/ MCNN
Inertia weight (W)	0.5	Randomization parameter (α_t)	1
Random number (r_1, r_2)	$0 \leq 1$	Attraction parameter (β_0)	0.97
Correlation factors (C_1, C_2)	0.5, 0.5	Absorption coefficient (γ)	0.01
Number of particles	10	No. of fireflies	10
Max. iteration	05	Max. iteration	5

Table 4 Values of evaluation metrics for CNN trained on different optimization technique

Training performance					Testing performance			
Evaluation metrics	RS-CNN	PSO-CNN	FA-CNN	FA-MCNN	RS-CNN	PSO-CNN	FA-CNN	FA-MCNN
MSE	1222.99	1045.31	1097.15	640.63	12,575.76	4313.72	2107.16	882.08
MAE	24.93	23.08	23.39	17.58	98.53	56.69	39.91	26.41
RMSE	34.97	32.33	33.12	25.31	112.14	65.67	45.90	29.69
MAPE	5.27	4.74	4.97	3.49	6.58	3.78	2.67	1.78

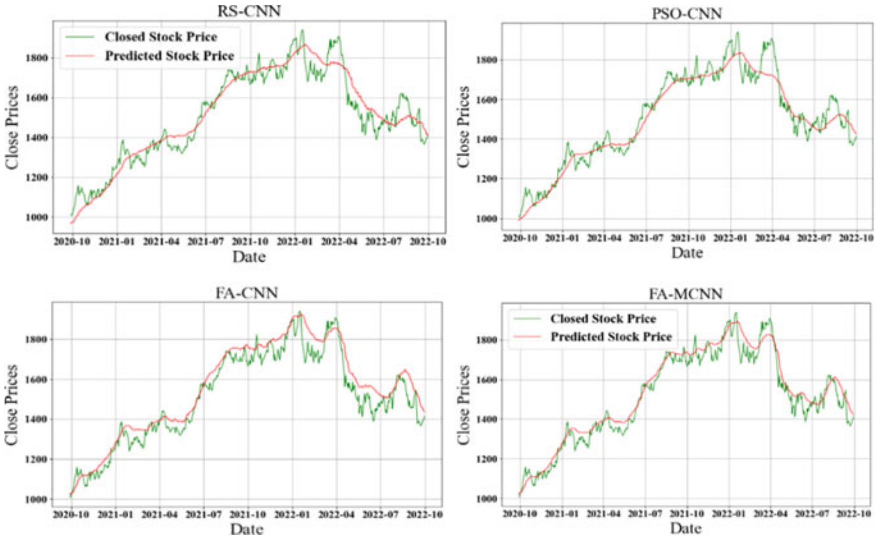


Fig. 4 Next day stock price forecasting performance

Since the parameters in the RS approach are selected at random, it is not a good strategy to identify the ideal parameter in comparison to other techniques. In contrast to PSO, which needs more iterations to discover the ideal parameter, the FA does so in fewer iterations. Because multiple CNNs are used in forecasting and the MCNN calculates the mean of all anticipated output by individual CNNs, its accuracy is higher than that of other techniques. The results show that the proposed FA-MCNN approach performs better than RS-CNN, PSO-CNN, and FA-CNN because it is more accurate at forecasting training and testing data. The forecasting performance of CNN trained on various optimization techniques for the training and testing datasets is presented in Figs. 4 and 5.

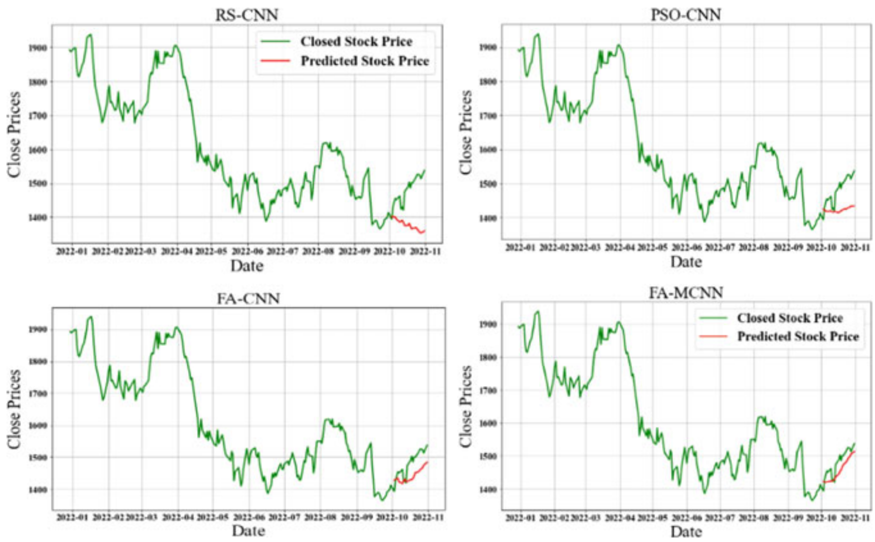


Fig. 5 October month stock price forecasting performance

5 Conclusion

In this study, we presented an MFA-MCNN model to enhance CNN forecasting accuracy and optimize CNN architecture parameters. The results show that PSO reduces costs with increasing particle size and epoch, but FA can do this with lower firefly counts and epoch. The MCNN calculates the mean of all anticipated output by individual CNNs, enabling forecasting with multiple CNNs more accurate than forecasting with other techniques. FA-MCNN can accurately predict stock values and offer pertinent information for investors to make a substantial profit.

References

1. Elangovan R, Irudayasamy F, Parayitam S (2022) Month-of-the-year effect: empirical evidence from Indian stock market. Asia-Pacifc Finan Markets 29:449–476. https://doi.org/10.1007/s10690-021-09356-2
2. Zargar F, Kumar D (2019) Opening noise in the Indian stock market: analysis at individual stock level. Theor Econ Letters 9:21–32. https://doi.org/10.4236/tel.2019.91003
3. Sayavong L, Wu Z, Chalita S (2019) Research on stock price prediction method based on convolutional neural network. In: 2019 international conference on virtual reality and intelligent systems (ICVRIS), pp 173–176. https://doi.org/10.1109/ICVRIS.2019.00050
4. Qolomany B, Maabreh M, Al-Fuqaha A, Gupta A, Benhaddou D (2017) Parameters optimization of deep learning models using particle swarm optimization. In: 13th international wireless communications and mobile computing conference (IWCMC), pp 1285–1290 (2017). https://doi.org/10.1109/IWCMC.2017.7986470

5. Bacanin N, Bezdan T, Venkatachalam K, Al Turjman F (2021) Optimized convolutional neural network by firefly algorithm for magnetic resonance image classification of glioma brain tumor grade. J Real-Time Image Proc 18:1085–1098. https://doi.org/10.1007/s11554-021-01106-x
6. Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. J Mach Learn Res 13:281–305
7. Chen J, Jiang J, Guo X, Tan L (2021) A self-adaptive CNN with PSO for bearing fault diagnosis. Syst Sci Control Eng 9(1):11–22. https://doi.org/10.1080/21642583.2020.1860153
8. Abbas F, Yasmin M, Fayyaz M, Elaziz MA, Lu S, El-Latif A (2021) Gender classification using proposed CNN-based model and ant colony optimization. Mathematics 9(19). https://doi.org/10.3390/math9192499
9. Zheng Y, Liu Q, Chen E, Ge Y, Zhao JL (2014) Time series classification using multi-channels deep convolutional neural networks. WAIM 2014 LNCS 8485, pp 298–310
10. Subasi A, Amir F, Bagedo K, Shams A, Sarirete A (2021) Stock market prediction using machine learning. Procedia Comput Sci 194:173–179. https://doi.org/10.1016/j.procs.2021.10.071
11. Shen J, Shafiq MO (2020) Short-term stock market price trend prediction using a comprehensive deep learning system. J Big Data 7(66). https://doi.org/10.1186/s40537-020-00333-6
12. Che Z, Purushotham S, Cho K, Sontag D, Liu Y (2018) Recurrent neural networks for multivariate time series with missing values. Sci Rep 8. https://doi.org/10.1038/s41598-018-24271-9
13. Cebeci YE (2019) A recurrent neural network model for weather forecasting. In: 2019 4th international conference on computer science and engineering (UBMK), pp 591–595. https://doi.org/10.1109/UBMK.2019.8907196
14. Hoseinzade E, Haratizadeh S (2019) CNNpred: CNN-based stock market prediction using a diverse set of variables. Expert Syst Appl 129:273–285. https://doi.org/10.1016/j.eswa.2019.03.029
15. Florea AC, Andonie R (2020) Weighted random search for hyperparameter optimization. Int J Comput Commun Control 14(2):154–169
16. Passricha V, Aggarwal RK (2019) PSO-based optimized CNN for Hindi ASR. Int J Speech Technol. https://doi.org/10.1007/s10772-019-09652-3
17. Fregoso J, Gonzalez CI, Martinez GE (2021) Optimization of convolutional neural networks architectures using PSO for sign language recognition. Axioms 10(3). https://doi.org/10.3390/axioms10030139
18. Shahzad A, Raza M, Shah JH, Sharif M, Nayak RS (2022) Categorizing white blood cells by utilizing deep features of proposed 4B-AdditionNet-based CNN network with ant colony optimization. Complex Intell Syst 8:3143–3159. https://doi.org/10.1007/s40747-021-00564-x
19. Chakraborty S, Aich S, Joo MI, Sain M, Kim HC (2019) A multichannel convolutional neural network architecture for the detection of the state of mind using physiological signals from wearable devices. Journal of Healthcare Eng. https://doi.org/10.1155/2019/5397814
20. Karhade J, Ghosh SK, Gajbhiye P, Tripathy RK, Acharya UR (2021) Multichannel multiscale two-stage convolutional neural network for the detection and localization of myocardial infarction using vectorcardiogram signal. Appl Sci 11(17). https://doi.org/10.3390/app11177920
21. Bao J, Li X, Gao Y (2019) MTCNN: a deep neural network for recognizing stochastic phase shifted time series data. In: 2019 Scientific Conference On Network, Power Systems And Computing (NPSC 2019), pp 201–206. https://doi.org/10.33969/EECS.V3.045

EDAMOD: Edge Detection and Moving Object Detection Algorithm to Control Traffic Lights Dynamically

S. Rakesh and Nagaratna P. Hegde

Abstract Recently, one of the biggest problems in most of the large cities throughout the entire world has been traffic congestion. The fast exponential growth of vehicles and the scarcity of appropriate roads to manage a large number of vehicles are the vital reasons for traffic congestion. Many researchers use frame differencing, edge detection (ED), and moving object detection (MOD) techniques individually to determine and predict traffic density; however, the ED method only detects edges for the input static image dataset, while the MOD paradigm determines the density of the traffic when vehicles are in motion. In reality, when a traffic intersection's red light is on and the cars are stationary, it is preferable to use the ED technique to estimate the traffic density lively. Conversely, when the green traffic signal comes on and the cars instantly begin to move, it is optimal to use the MOD approach. This study demonstrates a novel strategy known as the Edge Detection and Moving Object Detection (EDAMOD) approach, which partitions the Region of Interest (ROI) into ROI-1 and ROI-2 and then employs ED, MOD paradigms to identify the live area wide vehicle capacity at the traffic signal junction. In ROI-1, it employs ED, and in ROI-2, MOD.

Keywords Background subtraction · Edge detection · Vehicle density

S. Rakesh (✉)
IT Department, Chaitanya Bharathi Institute of Technology, Osmania University, Hyderabad, India
e-mail: srakesh_it@cbit.ac.in

N. P. Hegde
CSE Department, Vasavi College of Engineering, Hyderabad, India
e-mail: nagaratnaph@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_47

1 Introduction

Traffic management is one of the most significant and important issues in many large cities across the world. Poor traffic management can occasionally cause accidents that cause severe injuries or even fatalities. Traffic congestion is the main factor in the majority of automobile accidents that result in fatalities in cities. As more vehicles enter the congested traffic highways, new tactics must be implemented to route the limitations of the current traffic control conditions, because it is exceedingly expensive and time-consuming to construct new flyovers, elevated expressways, highways, etc. Therefore, one of the key objectives is to create a new traffic control strategy while employing present infrastructure and technological improvements.

The current world has a lot of digital and video data, and by processing that data using different technologies, a huge amount of information is produced. Closed-circuit television (CCTV) cameras had been deployed in various locations throughout all major cities, including the junctions of traffic light. Lots of video data is recorded by CCTV cameras and stored in different databases. This information is particularly beneficial in assessing and creating a solution for the traffic issues that most large cities in many nations experience. To comprehend and evaluate this enormous collection of data, a variety of image data analysis paradigms are tested, especially those that use the image data recorded and stored by the installed CCTV cameras arranged at traffic light junctions. Among the various image data analysis techniques, ED and MOD approaches are essential since they have many real-time applications.

2 Literature Survey

The approach now most frequently employed to address the issue of moving object recognition is the background subtraction method [1]. The creation of a background model forms the method's foundation. We obtain a backdrop image free of moving items once the background model is constructed. However, in some environments, getting the background can be challenging, and in some unique situations, such changing the lighting intensity, adding or removing things from the background, the background will always vary. As a result, the background representation model needs to be reliable and flexible. Different models can be created using the conventional background subtraction method depending on the modeling approach. Modeling the background using mean, median, or histogram analysis is the most fundamental model. After obtaining the background image, the difference threshold between the current image and the background image can be used to determine whether the pixels in the current image are foreground or background. Gaussian, support vector, subspace, and other models are often used difference threshold models. Lin et al. [2] suggested using probabilistic support vector machines to initialize the background for the purpose of using background subtraction to detect moving objects, and Tavakkoli et al. [3] suggested using support vector data description to identify foreground or background pixels in video sequences. In Oliver et al. [4], principal component analysis was used to rebuild the background in the image.

To find moving objects, a variety of methods are employed. Some of the methods used to find moving objects are recursive, while others use non-recursive methods. Approximation media filters, mixtures of Gaussians (MoG), and Kalman filters are recursive MOD techniques. The median filter, frame differencing, and linear predictive filter are non-recursive methods. All of these strategies' benefits and drawbacks were thoroughly discussed by Chandrasekhar et al. [5]. Bouwmans et al. [6] and Yilmaz et al. [7] both provided explanations of the important algorithm, known as MoG, which is used to detect the foreground.

Canny edge detection techniques offer very accurate results in image processing, and Canny and Prewitt ED approaches are becoming more popular [8, 9]. The Canny ED paradigm described in [10] is utilized to find the vehicle density, and the evaluated vehicle density is then used to control the intelligent traffic system. The borders of the vehicles that can be seen in the provided images are all that is provided. Actually, we need the road space that cars occupy. To avoid frequent collisions [11, 12], heavy traffic, and accidents, it is essential to build and employ a smart traffic light control and management system that smartly regulates and manages the traffic signal at the intersection.

3 Methodology

The steps in the EDAMOD approach for smart traffic management are as follows.

(A) Separate the ROI into two parts.
(B) Use edge detection to identify whether there are any vehicles in ROI-1.
(C) To determine the vehicle density in ROI-2, use the Moving Object Detection approach.
(D) Making decisions.

This paper develops and presents a revolutionary technique. Making decisions is a key element in this unique technique. The suggested innovative solution employs two techniques, namely moving object detection and edge detection. The ROI at the traffic signal junction was separated into two parts using the suggested method, as depicted in Fig. 1.

Canny edge detection was initially used to locate the vehicle edges in the first section. If this value exceeds a certain level, a green light traffic signal will be given in respective direction; otherwise, if there aren't many vehicles present, a red light signal will continue for a little while longer before moving to the next side region1 and applying Canny edge detection once more.

After five seconds, the moving vehicle density in ROI-2 was calculated if the Canny result was more than a predetermined threshold value, at which point a green light signal would be sent in that direction. Once more, a decision is made depending on the density of the moving cars; if it exceeds a certain threshold, the green light

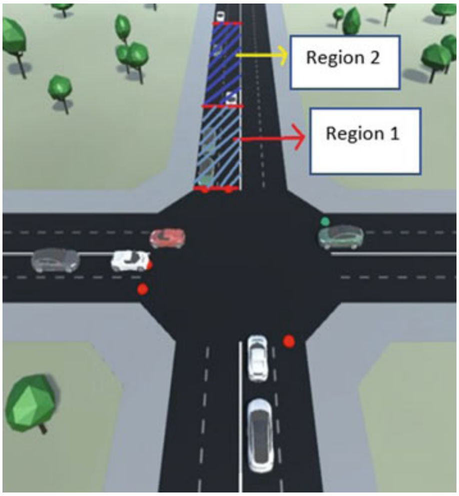


Fig. 1 Identifying ROI-1 and ROI-2 from total ROI

will remain on until it is depleted. Traffic lights must change from green to red with a 5-s grace period if the density of moving vehicles is below the threshold value, which indicates that there is very little traffic in that direction. This process resumed at the traffic light intersection in a loop. The flowchart in Fig. 2 presents the detailed controlled flow.

One of the very important steps in the proposed model is to find moving vehicles' density in ROI-2; for this purpose, background subtraction method mixture of the Gaussians algorithm is used. The steps included in finding vehicle density at this ROI-2 are shown in Fig. 3.

4 Results

In this section, the effectiveness of the suggested technique was evaluated with the aid of traffic surveillance recordings. The conversation between this examination and the evaluation is in-depth. All of these scenarios were tested using a Windows 10 system with an Intel Core i5 processor 2.5 GHz and 16 GB RAM. The processing time for this 320 by 240-pixel image ranged from 20 to 29 ms, and the sampling rate for the sequence was 30 frames.

Figures 4 and 5 illustrate sample frames that display the values for vehicle density as a result. Figure 4's 22 vehicle density indicates that the vehicles in the frame occupy 22% of the roadway. In a similar vein, Fig. 5's 19 vehicle density indicates that the vehicles visible in that frame occupy 19% of the road.

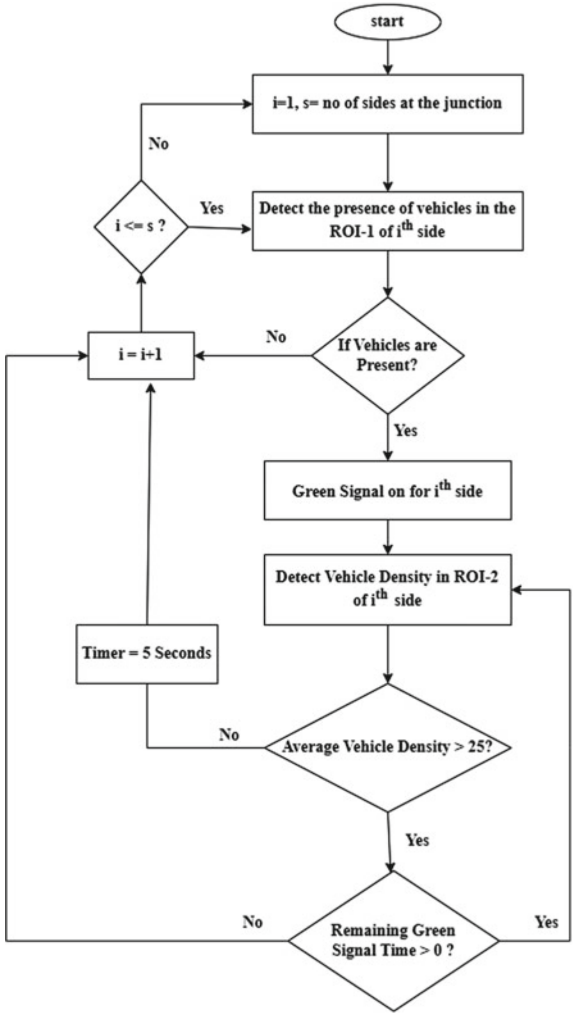


Fig. 2 Flow diagram of proposed model

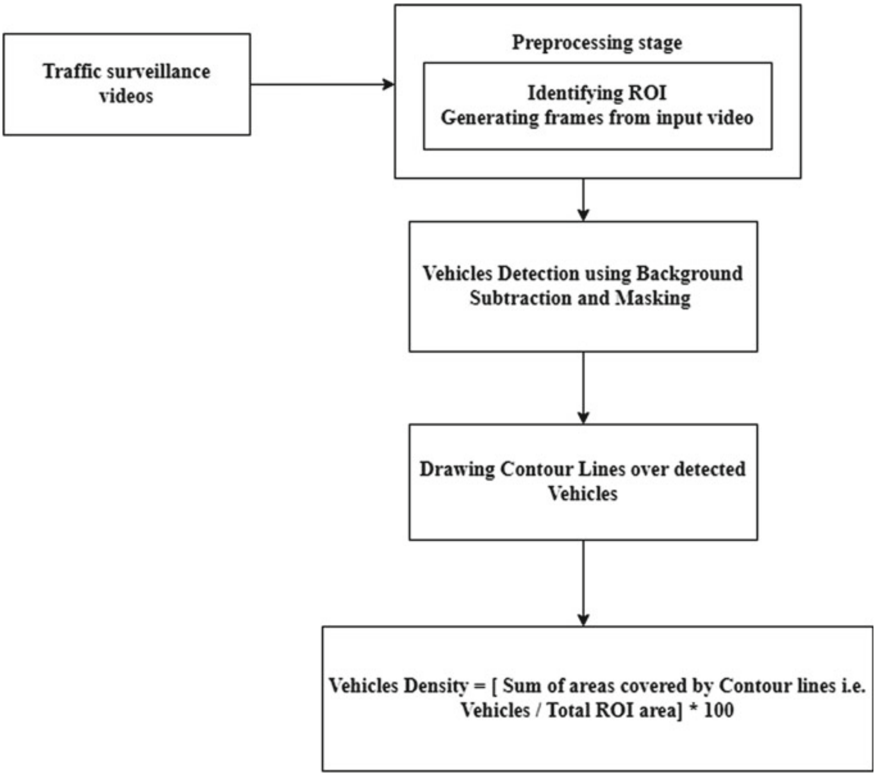


Fig. 3 Finding vehicles density using background subtraction

Fig. 4 Sample result displaying vehicle density value 22

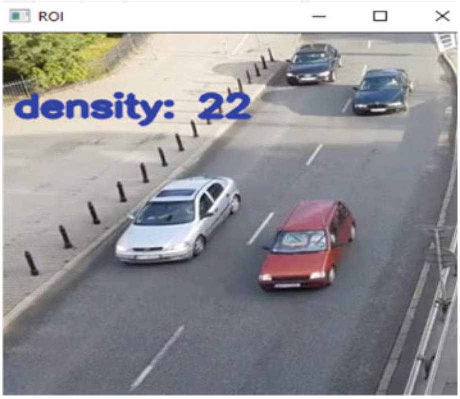


Fig. 5 Sample result displaying vehicle density value 19

5 Conclusion

The EDAMOD approach is an innovative methodology that operates effectively with both static frames and dynamic frames. It is proposed in this research as an EDAMOD model that works effectively and gives good results compared to previous strategies. By sending the live traffic videos into the suggested model as input, the proposed new approach effectively calculates the vehicle density values of live traffic dataset frames. The time needed to process and determine the traffic density for each frame is between 20 and 29 ms for the utilized traffic video datasets in this study, which can generate 30 frames per second.

References

1. Bouwmans T (2014) Traditional and recent approaches in background modeling for foreground detection: an overview. Comput Sci Rev 11:31–66
2. Lin HH, Liu TL, Chuang JH (2002) A probabilistic SVM approach for background scene initialization. In: Proceedings of international conference on image processing, vol 3. IEEE, pp 893–896
3. Tavakkoli A, Nicolescu M, Bebis G (2006) A novelty detection approach for foreground region detection in videos with quasi-stationary backgrounds. Advances in visual computing. Springer, Berlin, pp 40–49
4. Oliver NM, Rosario B, Pentland AP (2000) A Bayesian computer vision system for modeling human interactions. Pattern Anal Mach Intell IEEE Trans 22(8):831–843
5. Chandrasekhar U, Das T (2011) A survey of techniques for background subtraction and traffic analysis on surveillance video
6. Bouwmans T, El Baf F, Vachon B (2008) Background modeling using mixture of Gaussians for foreground detection—a survey. Recent Pat Comput Sci 1(3):219–237
7. Yilmaz A, Javed O, Shah M (2006) Object tracking: a survey. ACM Comput Surveys (CSUR) 38(4):13
8. Nisha RM, Lalita S (2015) Comparative analysis of Canny and Prewitt edge detection techniques used in image processing. Int J Eng Trends Technol (IJETT) 28(1)
9. Susmitha A, Mishra I, Sharma D, Wadhwa P, Dash L (2017) Implementation of Canny's edge detection technique for real world images. Int J Eng Trends Technol (IJETT) 48(4)

10. Uddin MS, Das AK, Taleb MA (2015) Real-time area based traffic density estimation by image processing for traffic signal control system: Bangladesh perspective. In: 2nd international conference on electrical engineering and information & communication technology (ICEEICT), 21–23 May 2015
11. Gupta S, Agarwal K, Venu Gopalachari M (2022) Smart home infrastructure with blockchain-based cloud IoT for secure and scalable user access. In: Proceedings of the international conference on paradigms of communication, computing and data sciences PCCDS 2021. Springer Nature
12. Reddy EA, Kavati I, Rao KS, Kiran Kumar G (2017) A secure railway crossing system using IoT. In: 2017 international conference of electronics, communication and aerospace technology (ICECA), vol 2. IEEE, pp 196–199

Student Monitoring System Using Deep Learning

D. Jayaram, S. Rakesh, M. Venu Gopalachari, Sumadhura Gaddam, B. Kranthi Kumar Reddy, and Pavan Sai Pulluri

Abstract The process of teaching and learning is frequently seen as the most important one in a school. However, manually taking attendance and keeping track of the pupils in a class can be laborious tasks that have a negative impact on the teaching and learning process. Additionally, disobedience on the part of the students such as not paying attention, using a phone, wandering around, and sneaking out of class might exacerbate the problem. An automated system will help the teachers to focus on the main syllabus and students to concentrate on studies. The results that we obtain from the system can be used by the faculty to analyze what number of students are really interested in the subject. If teachers can understand the negative attitudes of their students, they will be able to change the learning environment for the students in more logical ways. To overcome this situation, our aim is to develop an automated system that will mark the attendance and monitor the attentiveness of students. The system records the whole session and identifies whether the student is concentrating in the classroom or not, then reports to the faculties. To check what amount of time students attended the lecture with full attention or without any misbehavior, eye aspect ratio is used.

Keywords Behavior · Eye aspect ratio · Face detection · Machine learning

1 Introduction

Education plays a critical role in influencing future generations. Most of the time, especially in a large setting, it becomes difficult for a teacher to keep track of every student and determine their level of enthusiasm in learning. In these situations, instructors frequently lack knowledge of the best strategy to use when instructing pupils with various learning capacities. To create a real-time vision-based smart

D. Jayaram (✉) · S. Rakesh · M. Venu Gopalachari · S. Gaddam · B. Kranthi Kumar Reddy · P. S. Pulluri
IT Department, Chaitanya Bharathi Institute of Technology, Hyderabad, India
e-mail: djayaram_it@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_48

classroom that automatically tracks students' attention and provides in-the-moment feedback to the teacher, this proposed system uses AI. Our technology will be able to track a student's attendance in the room and assess their level of focus and emotion. With the aid of the mentioned system, instructors can more quickly and precisely assess the level of student attention, enabling them to conduct the lesson in a way that may improve academic performance. Field of education took various changes during the transition period of pandemic. One of the outcomes of this transition was enabling technology-based learning via Google Meet, Zoom, etc. Monitoring students in such environment is difficult due to lack of access to visual aids like gestures, body language of students, etc. A monitoring system like this will check drowsiness and evaluate student's attentiveness in virtual environments as well. If scaled to a larger group, the same can be used for a classroom as a whole and notify the instructor.

2 Literature Survey

Reference Paper [1] enlisted the methods of recognizing objects comprising facial features in various multimedia input forms such as images and videos. It depicted the use of Haar Cascade from OpenCV to distinguish relevant features from irrelevant ones. And uses MATLAB to analyze the image. It is a web application using Raspberry Pi 3 model can be used with Bluetooth and wireless LAN. Paper [2] authors proposed an application where data are extracted from web (Instagram), and images are collected using Beautiful Soup library. Haar cascade classifier is used to detect human face. Haar features helps to detect face features and AdaBoost is used for feature selection. In [3], authors proposed a method to find human eye blink using facial landmarks. EAR (eye aspect ratio) is calculated for both the eyes using a mathematical formula. The output of this formula helps to decide the state of eyes. In [4], authors proposed a system that finds out drowsiness of the driver with eye blinks and yarns. Eye aspect ratio and mouth aspect ratio are used to know about consciousness of driver. Alarm is set up to notify the driver. In [5], authors offered a conventional machine learning approach (algorithms) to determine the condition of the human eye. To extract features, these systems make use of hardware elements like sensors and cameras. The data are then classified using a variety of techniques, including machine learning (ML) methods like support vector machine (SVM), convolutional neural network, and hidden Markov model. Reference [6] consists of setting up the system so that the web camera is pointed directly at the kids and is connected to a laptop. The system gives attendance reports for each session as well as attention and emotion report specific to each student and group of students. The saved student photographs were utilized to identify students using the Haar Cascade algorithm. The student activities are classified into two categories, high attention and low attention. In [7], student features are recorded in each frame and analyzed based on a variety of behavior's involving head movement, tongue movement, and eye movement. Utilizing the ERT algorithm, captured photographs will be compared to training data to detect all student attributes. The mouth and eye movements are controlled by various eye aspect ratios (EAR) and mouth aspect ratios (MAR). In [8], each frame records the student's features, which are then analyzed based on a range of behavior's involving head, tongue, and eye movement. To identify all student traits, collected photographs will be compared to training data using the ERT algorithm. Different eye aspect ratios (EAR) and mouth aspect ratios (MAR) influence the mouth and eye movements. In [9], based on static photographs of the human face, this work proposes a simple algorithm for feature extraction, classification of seven different emotions, and real-time facial expression detection. This is done by training a multi-layer perceptron (MLP) neural network using the aforementioned algorithm. The input image is initially subjected to some preprocessing to localize and clip off faces to categorize human faces. In [10], authors enlisted methods of face detection using Haar features. For accurate results, Haar features slightly modified to separate Haar features. These features that include do not care in-between white and black area which are useful for few types of images that have improper face angle or bad lighting, etc. Using cascading algorithm, these features are applied on the input images. In [11], authors proposed an application where Haar cascading features algorithm is used to detect the face, and then in order to recognize face, eigenface (set of eigenvectors of human face) is used. Eigenvectors are unique for each face. During training, eigenvectors are calculated for each face. In real time when it receives input face, it calculates eigen vectors for it, and using Euclidian distance formula, nearest eigenvector from set is considered, and the identity of the face will be revealed. In [12], authors offered a conventional approach to track the pupil of the eye. Haar cascading features algorithm is used to find the pupil area. Trackbars are used to keep track the pupil. The entire captured area is considered as Cartesian plain and has its own boundaries. Haar features provide special dot feature in order to detect pupil.

3 Methodology

The methodology used in Ref. [1] begins with camera setting up with Raspberry Pi and placed in a position such that the driver's face is covered properly. Using VNC viewer, the camera-captured data are displayed on the screen, for this Raspberry Pi must be connected to VNC viewer software. Real-time data will be streaming on this viewer. All the data which is collected through camera are stored in Raspberry Pi storage. The data in the Raspberry Pi storage are sent to cloud. Here cloud is used for storing images. Cloud users have their own cloud name, API key, API secret, etc. Details like these are important because this information is needed in programming section to connect Raspberry Pi with cloud. Now data can be downloaded from cloud to analyze it using MATLAB. The discovered face was provided by the detector as Matrix (Mx4). MATLAB can be used to determine whether the condition is true if the Raspberry Pi camera was unable to detect any faces in the image. In reference paper [2], initially web data extraction takes place, and Beautiful Soup library is used to collect image type data from web; here data are collected from Instagram application in JSON form. Haar cascade classifier for human face detection uses Haar features to detect human face features like eyebrows using edge feature, etc. The image is converted into integral image to reduce computation cost. If we consider all possible features, then it would be like 160,000 features which is impossible. To overcome this, AdaBoost algorithm is used. AdaBoost picks best features from 160,000. Cascade can recognize human faces by employing the Haar function. If this system successfully completes all of the processes, it will be able to recognize the image of a human face. If it skips a step, no human faces were found in the image, according to the indicator. In reference paper [3], human facial landmarks are detected using Dlib. Frames are preprocessed (converted into gray scale). A total of 68 landmarks are shown, but we are only interested in landmarks of eyes. Six landmarks are used to represent an eye. In order to find blink, vertical height of an eye is important (here landmarks act as coordinate points to find height, width, etc.). Eye aspect ratio is calculated for both the eyes. As human beings have different-sized eyes, two readings are taken which includes reading of normal eye (active eye) and current state eye (may be opened or closed). Current state readings are compared to normal eye readings to determine state of eye (closed or opened) (Figs. 1, 2).

In reference paper [4], drowsiness is detected. A wide angle, 140-degree color camera mounted to the dashboard of the vehicle. In such a way that it captures driver's face. Here in this application Gaussian blur is applied so that unnecessary background noise is ignored. Frames are collected for use in subsequent procedures in groups of 30 after the driver's first EAR has been calculated and saved. If dynamic EAR is 20% smaller than initial EAR in more than 20 frames, compare the two and alert the driver. The MAR is used to determine drowsiness through yarns; if the value above the cutoff for more than 25 frames, drowsiness is assumed. In both situations, the alarm signals the motorist. Then, using Convex Hull, it determines the shape of

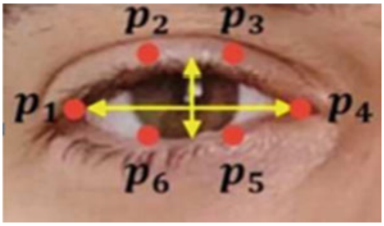


Fig. 1 Aspect ratio of eyes when open

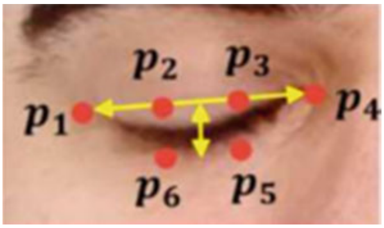


Fig. 2 Aspect ratio of eyes when closed

Fig. 3 Eye movements

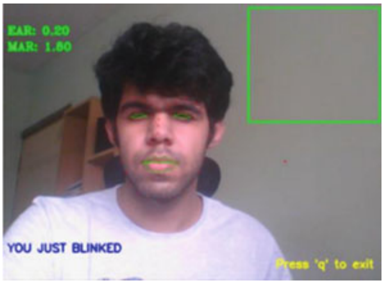
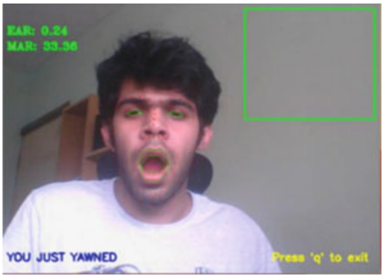


Fig. 4 Mouth movements



the object that was discovered in the ROI. When a hand is present, the alarm stops beeping and the driver's drowsiness is still being detected (Figs. 3, 4).

In reference paper [5], the project process takes place in three steps: (1) pretreatment, (2) feature extraction, and (3) model selection. Each frame of her BGR image is first turned into a grayscale image. The driver's face is then located in the frame using face detection, which makes use of the directional gradient histogram. To preserve the facial landmarks and extract the eye areas, use the shape predictor 68 face Landmarks DLIB module. His EAR is computed for each frame in step 3 and listed. The system starts sending his 15-element vector to the trained model in step 4 after storing the first 15 of his EAR values from the list. The input data are then divided into closed eye and open eye categories. If his eyes are closed for more than 15 consecutive times during the last step, it will stop the process. When a driver is deemed to be dozing off, an alarm goes off. In reference paper [6], this system captures student characteristics from each frame and examines that information in relation to different types of activities such as eye movements, mouth movements, and head movements to analyze student activity in each class. The captured images are compared with the data trained using the ERT algorithm to identify student features. A graph is used to show the overall performance of the students. DLIB's facial recognition prediction feature can detect drowsiness, yawning, and head posture. In reference paper [7], the deep learning model and a web camera make up the suggested system. The device continuously monitors the students during the lecture and plots in real time the status of their attention, emotions, and attendance. The camera captures a live feed from the class. The system is set up so that a web camera that is directly attached to a laptop faces students. The camera is mounted on the whiteboard and is set up to take pictures of the whole class. The system gives attendance reports for each session as well as attention and emotion report specific to each student and group of students. Yolo-v5 object detection method, used for its quicker and more accurate recognition, divides images into grids and cells. An object is tracked by the deep sort algorithm using detections over a bounding box. Every object in the frame is uniquely locked by the algorithm, which also identifies each one and tracks it until it leaves the frame. An essential component of this approach is the Kalman filter. The key benefit of the method is that it can track a single object without processing the entire video because it considers all the information in both the current and previous frames. Face recognition uses the Haar Cascade method. The stored student photographs allow for the identification of every student's face in the classroom. In reference paper [8], a drowsy driver can be identified using facial landmarks, DLIB, OpenCV, and Python. The I-Bug 300-W dataset [7] was used to train the localizer model for the DLIB, a pretrained facial landmark detector that can locate 68 face landmarks. These methods use facial landmark to track closed eyes and lips while identifying the position of the face, eyes, and lips. The goal is to extract the frames from the smartphone or web camera and extract facial landmarks, which translates into a set of x, y coordinates by identifying the locations of the facial landmarks. In this research, the device simultaneously monitors the openness or closure of the lips and the position of the eyes. The alarm starts sounding to warn the driver to avoid falling asleep behind the wheel whenever both lips are parted or both eyes are closed, which indicates that the driver is drowsy. The vehicle's engine will be stopped unless the driver regains consciousness.

4 Result Analysis

According to the reference study [1], the camera's position and angle are crucial for accurate findings and lowering the resolution will speed up the image-processing time. MATLAB is more precise than OpenCV at image capture and analysis (Haar Cascade). This is so because MATLAB offers a higher level of detail precision. Out of the 18 sample photographs, 6 can be recognized by MATLAB but not by OpenCV. Two of the remaining 12 samples can have mistaken according to MATLAB. MATLAB is 33.33% more accurate than OpenCV in this case. From reference paper [2], as the name suggests instead of applying all features at a time, it is done in stages. Only if the image satisfies stage n, then it goes to stage $n + 1$, if not it is classified as non-face. With this, only relevant image are filtered when we are trying to find Instagram pictures with the help of hashtag. From reference paper [3], finding eye blink using EAR is something more efficient than traditional machine learning classification. Here, it does not need huge amount of data to build model. Dynamically it detects state of the eye using eye aspect ratio. EAR formula is slightly modified for better performance irrespective of size of the eye. This paper is basic prerequisite to work on projects like drowsiness detection. From reference paper

[4], instead of independent on single EAR value per frame, it also calculates initial EAR value to compare it with real-time EAR value to know about state of the eyes. We do this to avoid inaccurate results caused due to different sizes of human eyes. There is no need to provide dataset because prediction does not depend on historic data, and by this, we can also avoid false alarms that cause due to false true. From reference paper [5], the ML classifier was used in the last stage to solve the false closed eye detection issue. An alarm is set off to alert the driver when sleepiness is found. Three ML models were tested after training stages, and the findings revealed that his RF model had the best performance with 99% accuracy. From reference paper [6], the graph will change or remain normal depending on whether the EAR is above or below the threshold. The yawn movement is counted once the MAR exceeds the threshold, and a graph is then drawn concurrently. It can tell whether a student is looking at the screen or not. The graph will increase quickly if he is not looking at the screen; else, it will be normal. From reference paper [7], there are several boundary boxes present all around the students. Multiple CSV files are used to store all the information offline. The information is retained in the sequence of discoveries. Thus, the information on the first student to be found will be saved first. Even after the lecture is complete, the instructor has access to all information. In online classes, the instructor's laptop screen displays each student's data during the lecture.

5 Conclusion

One of the important aspects of the project is drowsiness detection without training dataset because that might cause false alarms due to false predictions. EAR (eye aspect ratio) is perfect metrics to know about state of the eyes. To make it work for human beings of different eye sizes, it is better to compare real-time EAR with initial EAR (open state eye EAR). Following this technique, it will lead us to efficient results. Haar cascading features algorithm is the best one to detect the face, because it uses integral images which reduces computational time. And cascading algorithms are time saving as it is divided into stages even if one stage classified it as non-face then we can conclude it as non-face no need to go through rest of the stages.

References

1. Saad SL, Kamal MM, Zamri NA (2020) Monitoring and complaining web-based face recognition using Haar-MATLAB. IEEE
2. Priadana A, Habibi M (2021) Face detection using Haar cascades to filter selfie face image on Instagram. IEEE
3. Phuong TT, Hien LT, Toan DN, Vinh ND (2022) An eye blink detection technique in video surveillance based on eye aspect ratio. In: 2022 international conference on advanced communications technology (ICACT)

4. Chandiwala J, Agarwal S (2021) Driver's real-time drowsiness detection using adaptable eye aspect ratio and smart alarm. In: 2021 7th international conference on advanced computing & communication systems (ICACCS)
5. Redhaei AA, Albadawi Y, Mohamed S, Alnoman A (2022) Realtime driver drowsiness detection using machine learning. In: 2022 advances in science and engineering technology international conference (ASET)
6. Parambil MMA, Ali L, Gochoo M (2022) Smart classroom: a deep learning approach towards attention assessment through class behavior detection. IEEE
7. Chandra NBR, Suravarapu D, Goda RK (2022) Students live behaviour monitoring in online classes. IEEE
8. Al-madani AM, Shareef AAA, Ahmed ZAT (2021) Real-time driver drowsiness detection based on eye movement and yawning using facial landmark. IEEE
9. Haghpanah MA, Saeedizade E, Masouleh MT, Kalhor A (2022) Real-time facial expression recognition using facial landmarks and neural networks. IEEE
10. Jiang N, Lu Y, Tang S, Goto S (2010) Rapid face detection using a multi-mode cascade and separate Haar features. IEEE
11. Mantoro T, Ayu MA (2018) Multi-faces recognition process using Haar cascades and eigenface methods. IEEE
12. Li Y, Xu X, Mu N, Chen L (2016) Eye-gaze tracking system by Haar cascade classifier. IEEE

Sharing Logistic Information Files Securely Using Cryptographic Algorithms AES and SHA

Duvvada Rao Rajeswara Rao, Gopisetti PushpaLatha, Sakhamuri ShriVarshaa, and Vemula Vema Sri

Abstract As technology advances, data security has become a significant concern when using the internet. Because of rapid development, there has been a rapid change in the logistics industry, making logistics more informational, and allowing for the development of new business strategies. However, several issues must be addressed, such as ransomware and phishing attacks during file sharing or on unsecured remote desktops. As a result, to provide authenticity, confidentiality, and transparency to logistic information during file sharing, we propose a scheme using the advanced encryption standard (AES or Rijndael), an algorithm for symmetric encryption that maintains the confidentiality of logistic information, and SHA-256 for authenticating the data. Therefore, this project proposes an effective way to secure file transferring in cloud using AES and SHA algorithms.

Keywords Authenticity · Advanced encryption standard · Confidentiality · Logistic information · SHA-256

D. R. Rajeswara Rao · G. PushpaLatha · S. ShriVarshaa (✉) · V. Vema Sri
Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India
e-mail: shrivarshaasakhamuri@gmail.com

D. R. Rajeswara Rao
e-mail: rajeswararao.hod.cse@gmail.com

G. PushpaLatha
e-mail: gopisettipushpalatha7@gmail.com

V. Vema Sri
e-mail: vemasrivemula2001@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_49

1 Introduction

The development of technology over the past few years has resulted in the development of massive amounts of data almost instantly. In terms of real-time data processing and storage, this poses a challenge to the logistics industry, as this data is critical in developing new business strategies and identifying various patterns for the company's growth and development. However, many issues regarding the sharing of logistical information files must be resolved. To overcome this issue, we propose a web application that uses the AES and SHA-256 algorithms to share files in the cloud. Here, we are using AWS cloud service to store the files. This application can be run on any type of cloud. For data security, we are implementing this using the cryptographic algorithms AES and SHA-256. Throughout this case, we are encrypting the data with the conventional encryption algorithms AES or Rijndael.

We chose AES because it is a more secure, quick, and efficient cryptographic algorithm. Both hardware and software-sensitive data can be encrypted using AES. For administrative data protection, cybercrime, and the defense of electronic data, it is essential. The AES algorithm is widely regarded as uncrackable, at least not in our lifetime. AES provides three levels of encryption: 128, 192, and 256 bits. For 128-bit AES, encryption, and decryption, we use a 128-bit key size and 10 rounds. We use a key size of 12 rounds for encryption and decryption for the 192-bit AES algorithm and 14 rounds for encryption and decryption for such a 256-bit technique. We secure the data using the substitution, shift rows, mix columns, and addition of the round key operations, regardless of the key length. A 128-bit AES key, on the other hand, would take thousands of years to crack on a quantum computer. As a result, for data security, we propose using a 128-bit AES algorithm to encrypt data in the files. The SHA-256 algorithm is also used in this work because it is an excellent partner function to the AES algorithm. The SHA-256 algorithm succeeds the SHA-1 algorithm, which generates a unique 256-bit signature for the given data to provide hash codes to ensure that the data has not been tampered with. We are using the following combination of algorithms to achieve the data security and integrity of the shared logistic files.

2 Literature Survey

Chachapara et al. [1] presented private cloud functionality that would allow users to create keys for various users with various access levels to their files. Cryptographic algorithms such as AES and RSA are used in the framework. When a user creates a key, they have the option of giving it to the chosen user. As a result, the owner gives that user permission when they use that key to make a determined attempt to access files on the cloud. This is safer than giving the user a password and granting them only limited access. They have offered secure cloud sharing in this way. This technique enhances the safety of private cloud data by limiting access to a single user. It is easy and efficient to use. The drawbacks are, it limits the users, and for every user, we have to develop different keys. Sharing a group of files with the same key isn't secure.

Liang et al. [2] first enhanced the RSA key to a length where it could quickly produce large primes, thereby enabling the RSA algorithm in the cloud system. Hybrid encryption makes use of the RSA and AES algorithms. This hybrid encryption scheme protects lightweight data in cloud memory environments while also improving data confidentiality. This method ensures RSA encryption and decryption are quick. Inside a cloud-based storage system, it is practical, effective, and safe for users' portable data. The drawbacks are encryption and decryption becomes slower if it has more data. It can provide efficiency and security only to a certain extent.

Bansal et al. [3] proposed an RSA and Blowfish-based hybrid cryptosystem constructed in VHDL. Utilizing digital signatures for user authentication, this combination block cipher is used in cloud computing. This method offers symmetric and asymmetric cryptography features. This algorithm improves cloud computing security by being secure and enabling authentication. It provides security from brute force. The drawbacks are that it does not efficiently provide security as we can use devices with more resources and large key sizes.

Chen et al. [4] developed a straightforward AES, a hardware-based advanced encryption standard of FGPA. AES hardware structure design, extensive wiring with Quartus II 13.0, and simulation verification on Modelism SE10.5 are all included in the contract. This method ensures the high-speed AES designed by FGPA value. Fast and high real-time efficiency using AES. The drawbacks are high power consumption and expensive implementation. The implementation is more complex.

Malarvizhi et al. [5] proposed a system that enables mobile groups to archive and share data securely in the cloud. Any cloud user can share data effectively by using group signatures and data encryption. Without changing the remaining users' private keys, the user is revoked. Particularly for encrypting data files, the computation overhead is reduced, and the cipher text size is constant and independent of the revocation users. Due to the efficiency of the shared encryption and group signature, this method is more advantageous. The system enables newly registered users to instantly decrypt cloud-stored files belonging to other group members using their private key without first getting in touch with the file's owner. The disadvantages of this approach include storage overhead and the method takes longer to compute.

Zhao et al. [6] they proposed a privacy-preserving data search scheme capable of supporting both identifier-based and feature-based product searches two novel encrypted index trees are created, which can be searched without knowing the plaintext of the data. They created a system based on cloud computing for securely and effectively retrieving product information. This method ensures product information retrieval is secure and efficient. It ensures privacy through the use of encryption. The search process is extremely simple. The disadvantages include the difficulty in implementation. This scheme takes the longest to perform a search operation.

Ali et al. [7] they have developed a method to provide security for the shared data in the cloud. Nowadays, most organizations suffer from several security concerns. This system proposed a SeDaSC methodology for encrypting files with a single encryption key. Each user has two different keys, but only one of them is shared. The other key is shared with a third party that can identify insider threats, which is the Cryptographic server. In this way, the SeDaSC methodology provides confidentiality and integrity to the data that is shared.

3 Proposed Methodology

This section discusses the proposed system's methodology. The suggested system model for sharing logistic information files securely using cryptography techniques is depicted in Fig. 1. The proposed system clearly describes the entire process of what will happen in the application from start to finish in the manner specified.

3.1 Uploading and Encrypting the File

The AES encryption algorithm divides the data cached in it into blocks of 128 bits each. This encryption process employs the operations sub-bytes, shift rows, mix columns, and add round key.

Input: words in the file, i.e.,W_f [i,j] and key, i.e., K [i,j]
Output: cipher word of the file, W_c
We can generate W_f [i,j] as follows:
W_f = [[0 for x in range(4)] for x in range(4)]
block = range(128) //block size
For i = 1 to 4 do
  For j = 1 to 4 do
    W_f [j] [i] = block[32*i + 8*j:32*i + 8*(j + 1)]
Do 10 rounds for each byte in the array.

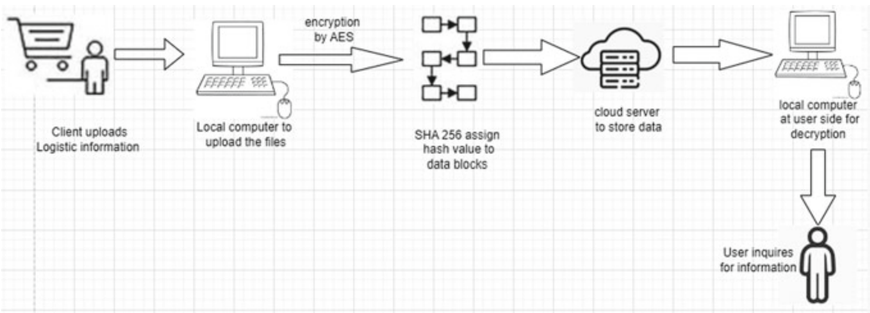


Fig. 1 Proposed methodology diagram

Step 1: Sub-bytes(for given byte)
Let b_{in} be the input byte and b_{out} be the output byte, then
For i = 1 to 4 do
For j = 1 to 4 do
  W^{SB} = [[0 for i in range(4)] for i in range(4)]
  W_f [i] [j] = b_{in}
  Function f ()
    $b' = b_{in} - 1$ //inverse
    $c = 0 \times 63$ //constant
Here, ‘A’ is 4 different circularly rotated versions of b′
    $b_{out} = (A \oplus b') \oplus c$
  $b_{out} = f(b_{in})$
 W^{SB} [i] [j]. append (b_{out})
Step 2: Shift rows(circular shift)
For i = 1 to 4 do
 For j = 1 to 4 do
  If (i = = 1) then
   W^{SR} [i,j] = W^{SB} [i, j]
  Else W^{SR} [i,j] = W^{SR} [i,j] << i
Step 3: Mix columns
W^{MC} [i,j] = W^{SR} [i,j] X C(z)
Where C(z) = Fixed polynomial coefficients
Step 4: Add round key
W_c [i,j] = W^{MC} [i,j] $\oplus$ k [i,j]

3.2 Using SHA-256 on the Encrypted Data

The data, i.e., stored in the cloud after the AES encryption is further given as input to the SHA-256 algorithm to generate hash values.

Input: The blocks of data in the file B_f
Output: 256-bit hash value for B_f
Step 1: Processing of the data
Convert B_f into binary code
B = binary(B_f)
B + = 1
B = B . append (0) until B_f is multiple of 512, less than 64 bits
B . append (len (B_f))

Step 2: Initializing 8 hash values—h0, h1, h2, h3, h4, h5, h6, h7, buffers—a, b, c, d, e, f, g, h with hexadecimal values and round constants, further used in the 64 rounds.

Creating 64(k[0.63]) constants where each number (063) corresponds to the initial 32 bits of the initial 64 prime numbers’ cube roots’ fractional bits (2–311).

Step 3: Processing the 512-bit blocks
For each block do
Initialize array of 32-bit w[0–63] words, copy the block into w[0–15] do
for i from 16 to 63 do
a := (w[i-15] right rotate 7) ⊕ (w[i-15]right rotate 18) ⊕ (w[i-15] >> 3)
B := (w[i-2] right rotate 17) ⊕ (w[i-2]right rotate 19) ⊕ (w[i-2] >> 10)
w[i] := w[i-16] + a + w[i-7] + b
Step 4: Compression function
For i from 0 to 63 do
r = (e right rotate 6) ⊕ (e right rotate 11) ⊕ (e right rotate 25)
ch := (e ∧ f) ⊕ ((!e) ∧ g)
t1 := h + r + ch + k[i] + w[i]
s := (a right rotate 2) ⊕ (a right rotate 13) ⊕ (a right rotate 22)
maj := (a ∧ b) ⊕ (a ∧ c) ⊕ (b ∧ c)
t2 := s + maj
h := g, g: = f, f: = e, e: = d + t1,
d := c, c: = b, b: = a, a: = t1 + t2

The compressed block is added to the current hash value—a, b, c, d, e, f, g, h values are added to the corresponding hash values h0, h1, h2, h3, h4, h5, h6, h7.

Step 5: The final hash value is.

Fhash = h0 append h1 append h2 append h3 append h4 append h5 append h6 append h7.

3.3 *Decrypt and Downloading the File*

Focuses on the decryption of AES and the sender sends the secret key to the receiver and which is used for downloading the file.

Decryption–AES : $W_f[i,j] = D_k AES(W_c)V$.

4 Results and Analysis

This section describes the system's results and output. Login page and sign up page and what functionalities the application provides (Figs. 2, 3, 4, 5, 6 and 7).

Fig. 2 Web application for sharing files

Fig. 3 Various options in the home

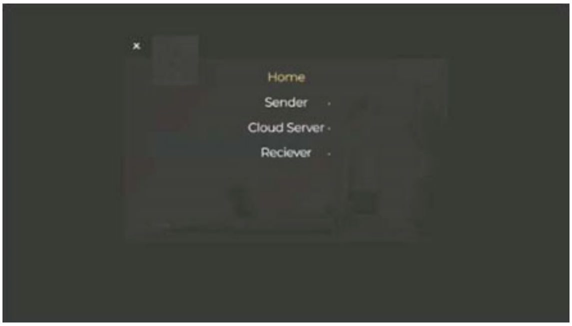
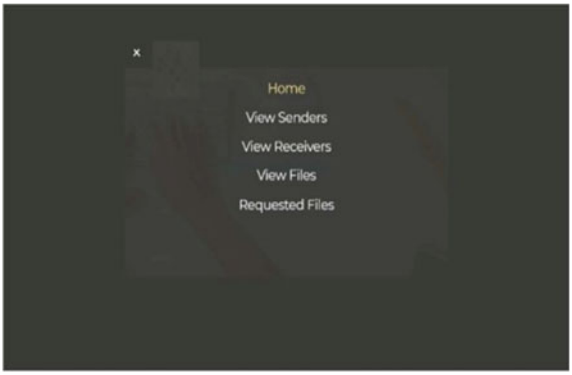


Fig. 4 Various options on the cloud



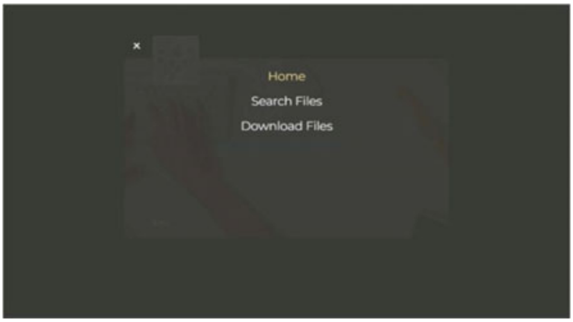


Fig. 5 Options on the receiver page

Fig. 6 Files to download by receiver

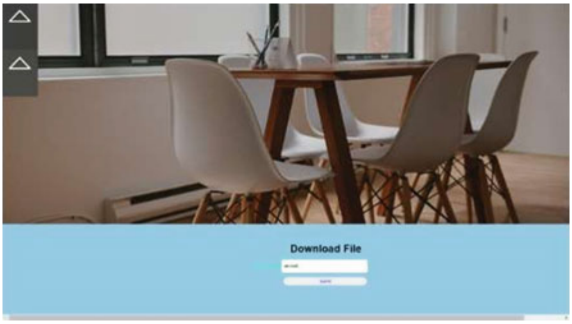


Fig. 7 Key to decrypt the file

5 Conclusion and Future Work

Under this paper, we propose a secure data exchange system and present the fundamentals of the AES and SHA-256 algorithms. Additionally, it covers how they are encrypted and decrypted. Data uploaders in the proposed system keep their data in the cloud in encrypted form. The authorized data requestors search the cloud for the necessary information. The search results serve as the basis for the requestors' request to the cloud for authorization to download data. Authenticated users can download and decrypt data using the key provided by the cloud. The suggested technique enables users to share logistics information securely, effectively, and confidentially over the cloud. For further work, we can add search retrieval of files, as of now we have included only a search of the file name for retrieval, we can further include the author of the file, etc. We can rank the properties such as security, authenticity and data integrity, and other factors and can further improve the model.

References

1. Chachapara K, Bhadlawala S (2013) Secure sharing with cryptography in cloud computing. In: IEEE 2013 Nirma University international conference on engineering (NUiCONE), 28–30 Nov 2013. https://doi.org/10.1109/TETC.2017.2737789
2. Liang C, Ye N, Malekian R, Wang R (2016) The hybrid encryption algorithm of lightweight data in cloud storage. In: IEEE 2016 ISAMSR, 23–24 Aug 2016. https://doi.org/10.1109/ISAMSR.2016.7810021
3. Bansal VP, Singh S (2015) A hybrid data encryption technique using RSA and Blowfish for cloud computing on FPGAs. In: IEEE 2015 RAECS, 21–22, Dec 2015. https://doi.org/10.1109/RAECS.2015.7453367
4. Chen M, Wei H, Li H (2020) Architecture design and hardware implementation of AES encryption algorithm. In: 2020 5th international conference on mechanical, control and computer engineering (ICMCCE), 2020, pp 1611–1614. https://doi.org/10.1109/ICMCCE51767.2020.00353
5. Malarvizhi M, Sujana JAJ, Revathi T (2019) Secure file sharing using cryptographic techniques in the cloud. In: 2019 international conference on green computing communication and electrical engineering (ICGCCEE), 2019, pp 1–6. https://doi.org/10.1109/ICGCCEE.2014.6921421
6. Zhao YS, Zeng Q-A (2018) Secure and efficient product information retrieval in cloud computing. IEEE Access 6:14747–14754. https://doi.org/10.1109/ACCESS.2018.2816919
7. Ali M et al (2017) SeDaSC: secure data sharing in clouds. IEEE Syst J 11(2):395–404. https://doi.org/10.1109/JSYST.2014.2379646

SVM-Based Vehicle Number Plate Detection and Recognition

S. M. Kazi and B. G. Kodge

Abstract The amount of automobiles has dramatically expanded during the last few decades. The number of user road cars has increased as a result of the rising wealth of modern living. For the sake of security, law enforcement, and traffic management, keeping track of every car is become exceedingly tiresome due to this constant increase. The need for a vehicle number plate recognition system is highlighted by the phenomenon of continually growing vehicle traffic. Numerous applications, including the upkeep of traffic laws, traffic control, automatic toll collection, parking systems, and automatic gate openers, utilise number plate detection systems. This work introduces a novel algorithmic approach based on the notion of mathematical morphology and SVM for the detection and identification of a car plate number. An SVM model is a representation of the examples as points in space, projected in such a way that the examples of the various categories are separated by a distinct gap that is as wide as possible. The created algorithm is straightforward, effective, and adaptable. The system also determines which Indian state the vehicle is from.

Keywords SVM · Image pre-processing · Number plate localisation

1 Introduction

VNPR systems are used to regulate security in restricted places like military campgrounds and protected sanctuaries, manage exit and access in car parks, collect toll payments, and more. These VNPR technologies are frequently used to bolster security in particular locations and prevent fraud. For instance, they may be useful when looking for stolen or crime-related automobiles. This task needs a significant amount

S. M. Kazi (✉)
Swami Ramanand Teerth Marathwada University, Nanded, Maharashtra, India
e-mail: kazism2020@gmail.com

B. G. Kodge
GITAM University, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_50

of labour, time, and resources, barring VNPR systems. Additionally, physical assistance in such jobs may result in incorrect interpretations, and in the meantime, it is almost impossible for a human to recall or read a moving vehicle's licence plate accurately.

The necessity of automatic licence plate recognition (VNPR) systems for automobiles makes this a difficult field of research for a variety of commercial applications. Localising the licence plate within a camera-captured image is the first and most crucial step in any VNPR system. This document includes the suggested methodology. It has taken into account Indian licence plates, whose rear adheres to licence plate standards. This system uses a several algorithms, including "feature-based number plate localisation" to locate the licence plate, the "image scissoring" technique to segment characters, and a suggested support vector machine algorithm to recognise characters (SVM). Single line licence plates are recognised by the system.

An image or video stream is often the input to a VNPR system, and if the provided frame contains a vehicle, the system outputs the content of the licence plate, typically as text. To take pictures of the automobiles, these devices have a camera. Depending on the needs of the system, those picture scans can be coloured, black and white, or infrared. The car number plate is detected and read using methods such image processing, object detection [1–4], and pattern recognition.

India's states and automobile codes are covered in Sect. 2, and a brief overview of related tasks is covered in Sect. 3. The proposed model is presented in Sect. 4, the experiment results are shown in Sect. 5, and the conclusion and future work are presented in Sect. 6.

2 Vehicle Codes for Indian States

Indian States Vehicle Codes and ISO 3166-2 for 28 States and 9 Union Territories such as MH-Maharashtra, AP-Andhra Pradesh, DL-Delhi. More than 23 crore vehicles were registered in India as of 2016, and all of the state-specific regional transport offices (RTOs) at the district level are responsible for issuing vehicle codes. The new BH series has been implemented from 15, September 2021, e.g. "21 BH 9999 UP" for a vehicle registered in 2021.

3 Related Work

It is difficult to offer a universal solution to detect and recognise a licence plate anywhere in the globe because different regions' licence plates different in size, colour, font, and standards. The applicability of the majority of the available methods to real-world applications is also limited by other considerations including licence plate rotations and occlusions. Because of this, solving VLPR under current restrictions requires the use of complicated and computationally costly approaches [5–7].

The extraction or detection of licence plate data is the initial step. To find the licence plate in a picture, existing algorithms combine object detection with computer vision techniques. The characteristics of the licence plate, such as its shape [2, 8], colour, symmetry, and texture are the basic foundation for computer vision algorithms.

The characters are recovered from the segmented licence plate in the second step utilising methods that are widely used, including mathematical morphology [9], connected components, relaxation labelling, and vertical and horizontal projection. However, not all multi-stage VLPR systems necessarily perform the character segmentation stage. The recognition of the characters using pattern matching methods or classifiers like SVM, neural networks, and fuzzy classifiers is the last stage.

Single-stage processes had recently undergone a number of successful tries. To the best of our knowledge, each of these attempts makes use of a single deep neural network that has been trained for end-to-end localisation, detection, and recognition of the licence plate in a single forward pass. Recognition of licence plates is a particular instance of object detection. These models can take use of the fact that licence plate detection and recognition are highly linked, just like single-stage object detectors [10]. In comparison with a standard two-stage model, this enables models to exchange parameters and have fewer parameters. They may therefore be quicker and more effective than a comparable two-stage technique [10, 11]. The table gives literature review which includes references to prior works that are pertinent and make use of different methodologies (Table 1).

Table 1 Review on licence plate detection and recognition

Reference number and year	Author	Technique
[12] 2021	Firasanti, T. E. Ramdhani, M. A. Bakri and E. A. Zaki Hamidi	OCR with Raspberry Pi
[13] 2021	L. Kong, Y. Bao, L. Cao and S. Zhao	Pyramid network
[14] 2021	W. Thumthong, P. Meesud and P. Jarupunphol	YOLO
[15] 2022	R. J. Tom, A. Kumar, S. B. Shaik, L. D. Isaac, V. Tripathi and P. Pareek	Deep neural network
[16] 2022	X. Fan and W. Zhao	Centre net
[17] 2022	A. Ashrafee, A. M. Khan, M. S. Irbaz and M. A. A. Nasim	MobileNet
[18] 2022	Mojtaba Shahidi Zandi Roozbeh et al.	CNN YOLO3

4 Methodology

SVMs have grown in significance in the field of pattern recognition since the 1960s. SVM is fiercely vying for classification supremacy against a variety of techniques. In order to get the best generalisation, SVM uses the statistical learning theory (SLT) as its theoretical basis and structural risk minimisation as its ideal object. They offer a clear intuition of what learning through examples is all about and are founded on a few basic notions. They have the quality of excellent performance in real-world applications, which is more significant.

The suggested system for vehicle number plate detection, recognition, and state detection (VNPR) has four stages: state detection, character segmentation, optical character recognition (OCR), and licence plate localisation (LPL). To find and pinpoint the location of a licence plate, LPL examines every pixel included in an image. The process of character segmentation involves identifying and separating each character on a licence plate. The state is determined via state detection. Various phases are as follows: image acquisition, pre-processing, plate detection, character segmentation, character recognition, state extraction.

Image acquisition—The image should be delivered in such a way that it shows the car's licence plate number and either the front or rear of the vehicle. System noise, blur, distortion, and other elements can all have an impact on how successful this stage is. *Image pre-processing* involves changing the image from RGB to grayscale. A two-dimensional grayscale version of the three-dimensional RGB image is created. *Number plate recognition*—Number plate recognition looks for the precise features in the image that contain the licence plate. *Character segmentation* is the technique used to cut out the necessary area of the licence plate. To correctly recognise each text, it requires separating the image from its background. *State extraction*—From the recognised alphanumeric characters, state is identified.

Given the right size and shape of a structural piece, morphology is a potent technique for classifying different elements within given images. The position, augmentation, and deletion of undesired image parts depend on the size and shape. Morphological operators have a crucial role to play in the extraction of a set's pertinent structures. That particular set interacts with a structure element to cause this procedure to happen. The morphological operator's dilatation and erosion make up the fundamental pair. The initial step in pre-processing for licence plate localisation (LPL) is to turn the original colour image into a grayscale image. The following phase involves scaling the image augmentation via morphology transformation, binary thresholding, the use of closing and opening morphology features to reduce noise, and licence plate extraction.

Adaptive median filter is applied to remove the noise from the image. Image processing technique such as edge detection, thresholding, resampling, and filtering have been used to locate and isolate the licence plate and the characters. The state is then detected through character extraction after character recognition using SVM. We made use of the Kaggle data set and Google Image.

5 Results

After the system receives the image, if it detects the licence plate, it displays the plate number along with the state or union territory to which it belongs; otherwise, if it doesn't (because of image noise, a tilted plate, excessive illumination, etc.), "licence plate not recognised" is shown.

The system can recognise single line number plates under widely varying illumination conditions with a success rate of about 90%. The experiment results so far are satisfying and promising (Figs. 1, 2, 3 and Tables 2, 3).

Fig. 1 Original image

Fig. 2 Localisation and segmentation

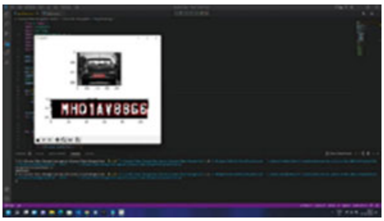


Fig. 3 State extraction

Table 2 Some result images

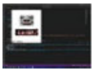
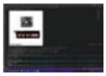
S. No.	Source image	Result image	S. No.	Source image	Result image
1			2		

Table 3 Result analysis

Applied stages	Accuracy (%)
Extraction of plate region	100
Segmentation	100
Recognition of alphanumeric characters	90
State or union territory identification	90

6 Conclusion and Future Scope

Several algorithms for finding licence plate information are currently available. Each method has an own set of limitations and applications. This work utilises an edge detection-based technique for positioning employing mathematical morphology to create licence plates. A combination of mathematical morphology and the edge detection filtering function is used to quickly place the licence plate region while reducing photo noise interference. The algorithm is rapid, dependable, and simple to use.

Future studies should concentrate on techniques for real-time licence plate detection and the extraction of intrinsic images (such as illumination, reflectance, and depth images) from movies. The licence plates can be used with licence plates from other nations even though they are designed specifically for use with Indian licence plates.

References

1. Montazzolli S, Jung C (2017) Real-time Brazilian license plate detection and recognition using deep convolutional neural networks. In: Proceedings of 30th SIBGRAPI conference on graphics, patterns images, Oct 2017, pp 55–62
2. Hsu G-S, Ambikapathi A, Chung S-L, Su C-P (2017) Robust license plate detection in the wild. In: Proceedings of 14th IEEE international conference on advanced video signal based Surveill (AVSS) (2017), pp 1–6
3. Xie L, Ahmad T, Jin L, Liu Y, Zhang S (2018) A new CNN-based method for multi-directional car license plate detection. IEEE Trans Intell Transp Syst 19(2):507–517
4. Laroca R, Severo E, Zanlorensi LA, Oliveira LS, Goncalves GR, Schwartz WR, Menotti D (2018) A robust real-time automatic license plate recognition based on the YOLO detector. In: Proceedings of international joint conference on neural network (IJCNN), July 2018, pp 1–10
5. Wang D, Tian Y, Geng W, Zhao L, Gong C (2020) LPR-Net: recognizing Chinese license plate in complex environments. Patt Recognit Lett 130:148–156

6. Rio-Alvarez A, de Andres-Suarez J, Gonzalez-Rodriguez M, Fernandez-Lanvin D, Pérez BL (2019) Effects of challenging weather and illumination on learning-based license plate detection in non-controlled environments. Sci Program 2019:1–16
7. Raghunandan KS, Shivakumara P, Jalab HA, Ibrahim RW, Kumar GH, Pal U, Lu T (2018) Riesz fractional based model for enhancing license plate detection and recognition. IEEE Trans Circuits Syst Video Technol 28(9):2276–2288
8. Luo L, Sun H, Zhou W, Luo L (2021) An efficient method of license plate location. In: Proceedings of 1st international conference information science engineering, 2009, vol 9, pp 770–773
9. Zou L, Zhao M, Gao Z, Cao M, Jia H, Pei M (2018) License plate detection with shallow and deep CNNs in complex environments. Complexity 2018:1–6
10. Li H, Wang P, Shen C (2019) Toward end-to-end car license plate detection and recognition with deep neural networks. IEEE Trans Intell Transp Syst 20(3):1126–1136
11. Xu Z, Yang W, Meng A, Lu N, Huang H, Ying C, Huang L (2018) Towards end-to-end license plate detection and recognition: a large dataset and baseline. In: Proceedings of European conference on computing visual. Lecture notes in computer science, vol 11217, pp 261–277
12. Firasanti A, Ramadhani TE, Bakri MA, Zaki Hamidi EA (2021) License plate detection using OCR method with Raspberry Pi. In: 2021 15th international conference on telecommunication systems, services, and applications (TSSA), pp 1–5. https://doi.org/10.1109/TSSA52866.2021.9768252
13. Kong L, Bao Y, Cao L, Zhao S (2021)A key point-based license plate detection with pyramid network structure. In: 2021 2nd information communication technologies conference (ICTC), pp 249–253. https://doi.org/10.1109/ICTC51749.2021.9441496.Ascar [Davix Jadson D. et al. Segnet approach for license plate localization seventh International conference on information technology trends 25–26 (2020)]
14. Thumthong W, Meesud P, Jarupunphol P (2021)Automatic detection and recognition of Thai vehicle license plate from CCTV images. In: 2021 13th international conference on information technology and electrical engineering (ICITEE), pp 143–146. https://doi.org/10.1109/ICITEE53064.2021.9611826
15. Fan X, Zhao W (2022) Improving robustness of license plates automatic recognition in natural scenes. IEEE Trans Intell Transp Syst 23(10):18845–18854. https://doi.org/10.1109/TITS.2022.3151475
16. Tom RJ, Kumar A, Shaik SB, Isaac LD, Tripathi V, Pareek P (2022) Car license plate detection and recognition using modified U-net deep learning model. In: 2022 8th international conference on smart structures and systems (ICSSS), pp 01–06. https://doi.org/10.1109/ICSSS54381.2022.9782176
17. Ashrafee A, Khan AM, Irbaz MS, Nasim MAA (2022)Real-time Bangla license plate recognition system for low resource video-based applications. In: 2022 IEEE/CVF winter conference on applications of computer vision workshops (WACVW), pp 479–488. https://doi.org/10.1109/WACVW54805.2022.00054
18. Shahidi Zandi M, Rajabi R (2022) Deep learning based framework for Iranian license plate detection and recognition. Multimedia Tools Appl 81:15841–15858

Enhanced Adaptive Multi-level Clustering of Long Tail Items

Soanpet Sree Lakshmi, T. Adilakshmi, and Bakshi Abhinith

Abstract The interest in large-scale recommendation systems is growing owing to the tremendous increase in data capturing capacity and need to analyze the data thus captured. The recent explosion of available data is also triggered by the application of IoT to various fields. The need for studying the Large-Scale RS is all the more important with this kind of technology capturing data at a very high speeds, from variety of devices in various different forms. Most of the research in this field is toward handling the large data. We propose an adaptive multilevel clustering of Long Tail Items to handle the large data. The proposed idea is improving the recommendation accuracy through adaptive approach for large data with long tail items.

Keywords Long Tail problem · Adaptive Clustering · Large scale data · Recommender systems

1 Introduction

Long Tail problem of the recommendation systems is studied extensively in the literature [1, 2] and the benefit it offers to promotion of niche items in the commercial systems is well known [3]. Various approaches are proposed to address the problem of recommender systems. Adaptive methods are known to improve the quality of recommendations made [4]. The interest in large-scale recommendation systems is growing owing to the tremendous increase in data capturing capacity and need to analyze the data thus captured. The recent explosion of available data is also triggered by the application of IoT to various fields. The need for studying the Large-Scale RS is all the more important with this kind of technology capturing data at a very

S. S. Lakshmi (✉)
Department of Information Technology, Vasavi College of Engineering, Hyderabad, India
e-mail: s.sreelakshmi@staff.vce.ac.in

T. Adilakshmi · B. Abhinith
Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_51

high speeds, from variety of devices in various different forms. Most of the research in this field is toward handling the large data. We propose an adaptive multilevel clustering of Long Tail Items.

The use of multilevel clustering techniques is studied in literature for use in large-scale RS. The multilevel structure of clusters can have two or more levels. Basic advantages of this multilevel clustering are that there can be multiple ways of analyzing the items. For example, the movies may first be clustered based on movie profiles, then in the next level by demographics, and then next level by language and so on.

2 Related Work

Authors in [6] used this multilevel cluster analysis to patients visit to doctor in studying the patient level, physician level and practice-level factors related to visiting the doctor.

Authors in [7], solved the problem by hierarchically clustering the items through different levels from coarse level to the finer levels. The authors demonstrated the effectiveness of this hierarchical method to single level clustering.

k-means [5] is a simple and well-known algorithm used to solve clustering problems. The goal of the algorithm is to find the optimal division of n objects into k clusters so that the total distance between the members of the cluster and the corresponding centroids representing the cluster is minimized. This algorithm uses an iterative improvement strategy. Following steps explain *k*-means methodology.

In this step, randomly k centroids are selected. These points act as seed points which will further be refined.

In this step, every object is given a clusterid depending on the closest centroid. The closeness is calculated using the Euclidean distance measure. The distance of each point to all the centroids is calculated.

The centroid values are recomputed and updated by average values of point dimensions/object features.

The above two steps are repeated many times till the cluster composition does not change in successive iterations.

Thus, the large-scale RS can be handled using this multi-level approach.

Authors in [8] provided a comprehensive survey of techniques of Recommendation Systems. They presented the focus on various studies in the fields of RS.

Authors in [9] presented the challenges for large-scale social networks. The challenges of data volatility, variety, volume is discussed in the chapter. Authors presented some of the solutions used in the context to overcome such challenges. Use of distributed computing over big data is the main focus of the paper.

Authors in [10] proposed a scalable approach to solve recommendation system for a large-scale RS. They proposed construction of large graphs used partitioning methods to handle scalability.

We propose to study the impact of Multilevel clustering for EADCCC-LT on accuracy of the long tail recommendation. This aspect is explained in the next section.

3 Multilevel Clustering of Correlation Connected Clusters

Figure 1 displays a multilevel data structure with three levels, where movies are clustered in Level 1 into five clusters. These five clusters are then clustered in the Level 2 into two clusters.

The multilevel clustering helps to minimize the total number of cluster search operations.

In the Proposed work, the movies are first clustered based on CCC method then they are further clustered based using *k*-means method. The optimal number of clusters is found using the elbow method.

3.1 Algorithm: EADCCC-LT-MuL

The method Proposed here is two level clustering. The correlated connected clusters generated at the lowest level are clustered further using *k*-means in the second level. This summarization of clusters at more levels helps reduce the recommendation effort to a greater extent (Fig. 2).

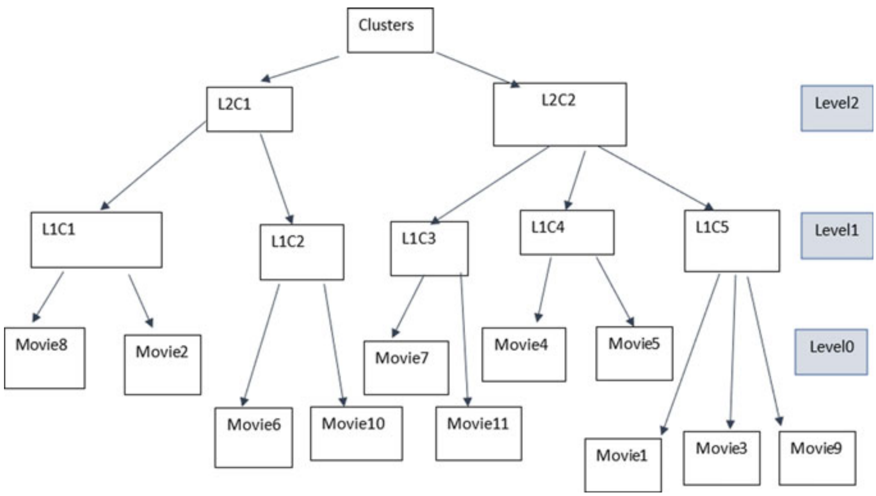


Fig. 1 Shows a multilevel data structure with three levels

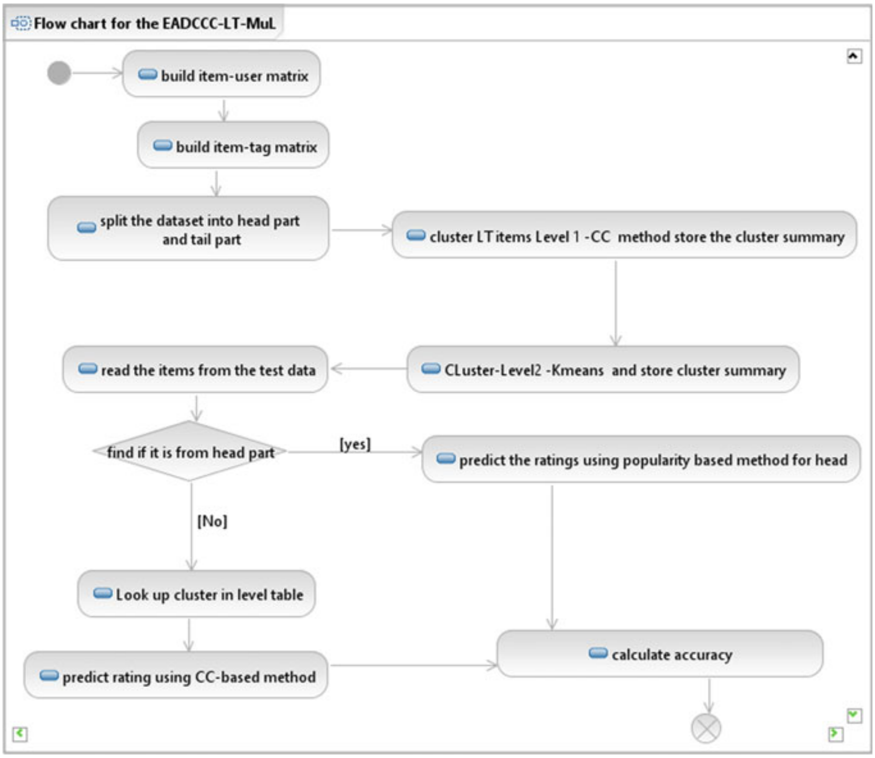


Fig. 2 Flow chart for proposed EADCCC-LT-MuL

Algorithm: EADCCC-LT-MuL

Step1: Calculate variables related to user and movies:- User_favourite_genre, User_genre_average, Movie_average_rating, Movie_counts, Moid_genre

Step 2: Depending on movie_counts, the data is split into two parts: head and tail. The criterion that is used to split the data is alpha (α). It is calculated by plotting the sorted no of ratings and finding the elbow point of the curve.

(a) If the movie_counts > (α), step 3 is chosen.
(b) else if movie_counts <= (α) step 4 is chosen.

Step 3: Rating is predicted based on User_genre_average that is calculated for every pair of userid and movieid. (Movie with large number of ratings)

Step 4: Here the rating is predicted using the clusters that were generated in long tail.

(a) Correlation connected clusters are generated for the movies in long tail
(b) The cluster means for these generated clusters are found and are then clustered again using kmeans method – level 2 clusters

User_id	Movie_id	predicted rating	actual rating	level 1 cluster	level 2 cluster		squared difference	mean	RMSE
116	4678	2.55237981	2	1	0		0.305123459	0.835753	0.914195
135	4678	2.84585408	2.5	1	0		0.119615044		
271	4678	3.01767593	4	1	0		0.964960573		
347	4678	3.06052292	2	1	0		1.124708869		
415	4678	3.25506744	2	1	0		1.575194282		
436	4678	3.64429158	5	1	0		1.837945322		
549	4678	3.37200897	4	1	0		0.394372733		
572	4678	3.17469551	3.5	1	0		0.105823012		
648	4678	3.16694484	4	1	0		0.693980907		
702	4678	2.80284836	4	1	0		1.433172046		
775	4678	3.17357666	4	1	0		0.682975538		
792	4678	3.41190522	4.5	1	0		1.18395026		

Fig. 3 EADCCC-LT-MuL-RMSE value

(c) The means for these level2 clusters are also calculated
(d) Firstly, the level 2 cluster to which the movie belongs is found.
(e) Then, the movie is compared to the cluster means of level 1 clusters present in level 2 cluster.
(f) The level 1 cluster to which the movie belongs to is found.
(g) The rating is predicted as the average of user_genre_average and average ratings of all movies in the level 1 cluster that are of the same genre as that of movid_genre.

Step 5: The error between the actual rating and predicted rating is calculated.
Step 6: Accuracy is calculated for their predictions.
Step 7: Calculate error %

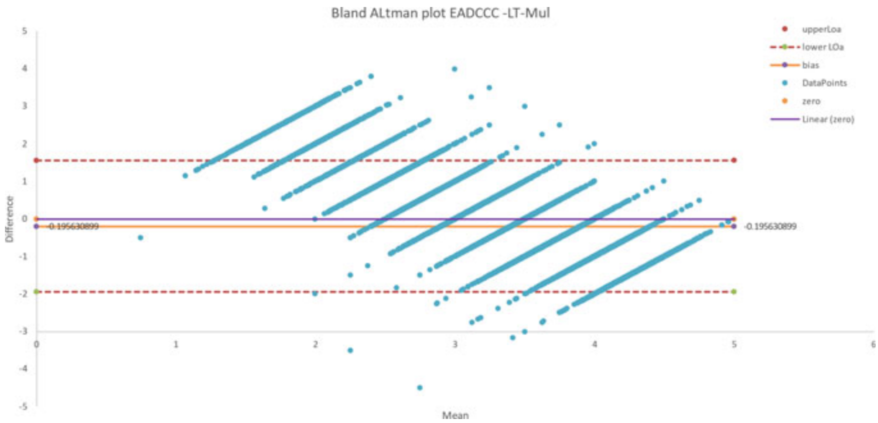
3.2 *Experimental Evaluation*

The RMSE is calculated for the test data as shown in Fig. 3.

The RMSE calculated for the test data is 0.91. This value is observed to be lower than that of the standard algorithms. The analysis of the data is performed and is shown in the next section.

3.3 *Result Analysis*

Inference: Our proposed model is underestimating the actual model with a Bias value of 0.19 that is the difference between the Zero Line of equality and our Bias line. Hence the model may be accepted and considered. In General, any system that generates a bias line between the Upper LOA and Lower LOA can be considered and accepted (Fig. 4).



	bias	**-0.19563**
	Std-dev	**0.893018**
bias-1.96*SD	**lower LOA**	**-1.94595**
bias+1.96*SD	**upper LOA**	**1.554684**

x-axis	**upper LOA bias+1.96*SD**	**lower LOA (bias-1.96*SD)**	**bias**
0	**1.554684407**	**-1.945946204**	**-0.19563**
5	**1.554684407**	**-1.945946204**	**-0.19563**

Fig. 4 Bland–Altman plot EADCCC-LT-Mul for 1000-users test data

4 Performance

See Table 1.

Table 1 Performance—EADCCC-LT-MuL

Model	RMSE
Proposed approach EADCCC-LT-MuL	0.914

5 Summary

Method	Clustering method
Adaptive clustering [4]	Ratings based
Proposed method EADCCC-LT-Mul	Multi-level clustering

The long tail items were clustered based on the correlation connected clusters concept in the first level thus enabling to find many arbitrary shaped correlated connected clusters. These clusters are further clustered in the second level using *k*-means algorithm. Applying the multilevel clustering concept, the results found to be similar to that of EADCCC-LT. The proposed clustering method used in long tail greatly reduced the error as shown by the RMSE values. The number of searches were greatly reduced in this approach by using the multilevel table structure.

References

1. Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: a survey of the state -of-the art and possible extensions. IEEE Trans Knowl Data Eng 17(6):734–749
2. Anderson C (2006) The long tail. Hyperion Press
3. Park YJ, Tuzhilin A (2008) The long tail of the recommender systems and how to leverage it. In: Proceedings of ACM conference on recommender systems, pp 11–18
4. Park YJ (2013) The adaptive clustering method for the long tail problem of recommender systems. IEEE Trans Knowl Data Eng 25(8)
5. Han and Kamber, Data mining techniques
6. Holodinsky JK, Austin PC, Williamson TS (2020) An introduction to clustered data and multilevel analyses. Family Practice 37(5):719–722. https://doi.org/10.1093/fampra/cmaa017
7. Bouhmala N (2016) A multilevel k-means algorithm for the clustering problem conference paper. https://doi.org/10.1109/ICCCBDA.2016.7529544
8. Wang M, Fu W, He X, Hao S, Wu X (2020) A survey on large-scale machine learning. IEEE Trans Knowl Data Eng. https://doi.org/10.1109/TKDE.2020.3015777
9. Eirinaki M, Gao J, Varlamis I, Tserpes K (2018) Recommender systems for large-scale social networks: a review of challenges and solutions. Future Generation Computer Systems 78, Part 1, pp 413–418. ISSN 0167-739X. https://doi.org/10.1016/j.future.2017.09.015
10. Bathla G et al (2020) Scalable recommendation using large scale graph partitioning with pregel and giraph. IJCINI 14(4):42–61. https://doi.org/10.4018/IJCINI.2020100103

UDP Flood DDoS Attack Detection Using Multilayer Perceptron Different Optimization Techniques with Pearson, Spearman and Kendall Uncorrelated Feature Subsets

Kishorebabu Dasari and Srinivas Mekala

Abstract The main challenge of the Internet is security. Distributed Denial of Service (DDoS) attacks is a major Internet security threat nowadays. Traditional detection methods are difficult to detect DDoS attacks because attackers use legitimate packets and frequently change package information. This research proposed a detection methodology with multilayer perceptron classification algorithm using feature selection correlation methods. And it is a quantitative study comparing a Multilayer Perceptron (MLP) classification algorithm ADAM, SGD, and LBFGS optimal approaches with distinct uncorrelated feature subsets for detecting UDP-flood DDoS attacks. Pearson, Spearman, and Kendall correlation methods are used to determine features. UDP-flood attack dataset collected from the Canadian CIC-DDoS2019 evaluation datasets. Experiment results conclude that all optimization techniques give more equal better results with all uncorrelation feature subsets.

Keywords UDP-flood DDoS attack · MLP classification algorithm · Correlation methods

1 Introduction

Distributed Denial of Service (DDoS) attacks [1] are a severe cyber attack that affects millions of websites every day. A DDoS attack is when one or more entities attempt to take down a website, server, or other internet-connected equipment by flooding it with

K. Dasari (✉)
Department of CSE (AI & ML), Keshav Memorial Institute of Technology, Hyderabad, Telangana 500029, India
e-mail: dasari2kishore@gmail.com

S. Mekala
Department of CSE, Keshav Memorial Institute of Technology, Hyderabad, Telangana 500029, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_52

huge amounts of traffic. The hosting servers are overloaded and unable to process all of the traffic, resulting in the site crashing or failing to respond to legitimate requests. These attacks are carried out for fun by hackers and other malicious actors who are paid by third parties, are politically motivated and are primarily intended to create a distraction.

Generally DDoS attacks classified into Volumetric, Protocol and Application layer based DDoS attacks. With huge amounts of malicious traffic, volumetric DDoS attacks are meant to exceed internal network capacity and even centralized DDoS mitigation scrubbing facilities. The first stage of a DDoS attack is to consume bandwidth within or between the target network/service and the rest of the Internet. UDP flood, Syn flood, and ICMP flood attacks are all common volumetric DDoS attacks. Protocol exploits target network layer and transport layer protocol interactions with malicious connection requests in order to use the processing capacity of network infrastructure resources such as servers, firewalls, and load balancers. A Smurf DDoS attack is a Common Protocol-based DDoS attack. Application layer DDoS attacks target certain applications, the most popular of which being web servers. Disrupting the servers by sending huge requests. Attacks on the application layer use fewer resources than volume-based and protocol-based attacks. Common Application layer DDoS attacks are Slowloris, Slow post, HTTPS flooding.

A UDP flood [2] is a volumetric DDoS attack in which the attacker floods the victim's computer with IP packets carrying User Datagram Protocol (UDP) packets. Massive UDP packets are sent to a specific destination or to random ports by the attacker. Because the UDP protocol is "connectionless," attackers can easily spoof the originating IP address. The fundamental purpose of a UDP flood is to saturate the Internet pipe. Another consequence of this attack is on network and security elements, particularly firewalls, along the path to the target server. Each UDP packet opens a state in a firewall, and the rush of connections quickly overwhelms them.

This section covers the DDoS attack and the UDP-flood DDoS attack. The proposed methodology is described in detail in Sect. 2 of this study. The experimental results are explained in Sect. 3. This research's conclusion and feature enhancement are covered in Sect. 4.

2 Methodology

The UDP-flood DDoS attack dataset collected from CIC-DDoS2019 [3] datasets and operations performed on it. Because the quality of data and the significant information that may be extracted from dataset has a direct impact on our model's capacity to train, so data preprocessing [4] is a pivot point in machine learning. Preprocessing process, remove constant socket features and remove missing and infinite value records. Use 0 and 1 to encode the Benign and UDP DDoS attacks, respectively. To improve the performance of the classification algorithms, standardize the feature values.

Feature selection [5] is a method of removing redundant, irrelevant, or noisy features from total features in order to pick a subset of the most relevant features.

Filter, Wrapper, Embedded, and Hybrid approaches are the types of feature selection methods. Using multiple metrics, the filter technique removes undesired features of the model. The feature selection in wrapper techniques is based on a machine learning algorithm that tries to fit a certain dataset. Wrapper and filter methods are used in embedded methods, which enhance feature interactions while maintaining the computational expenses low. In this study, feature selection is done utilizing filter-based feature selection methods such variance threshold and correlation methods [6]. All features whose variance fall below a certain threshold are removed b variance threshold. By default, it removes all zero-variance features that have the same value across all samples. It removes all 0.01-variance features, those features are called quasi-constant features. Correlation [7] is a statistical word that describes how closely two features are related in a linear manner. High correlation features have a similar influence on the dependent variable since they are more linearly dependent. When there is a strong association between two features, one of them may be dropped. The correlation coefficients' values range from − 1 to + 1, indicating how strong the relationship between the features is. When the coefficient value is ± 1 it represents high correlation among features. When the coefficient value is zero, the traits are highly unrelated. The Pearson, Spearman, and Kendall [3] correlation approaches are used in this study to identify uncorrelation properties. A linear relationship between two continuous variables is measured using Pearson's. Spearman's is a nonparametric test that uses ranked data to measure a monotonic connection. Spearman's has the advantage of being simpler to calculate. Kendall's method is nonparametric, which means it does not require the two variables to be on a bell curve. Kendall's does not require continuous data as well. It will work with continuous data because it is based on the ranked values of each variable, but it can also be used with ordinal data. Although ordinal data have a ranking, the intervals between rankings are not always consistent.

The correlation coefficient calculated by Pearson is

$$r = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum(X_i - \overline{X})^2 \sum(Y_i - \overline{Y})^2}} \tag{1}$$

Here r is the correlation coefficient

X_i is the X-feature values in a sample.

$\overline{X}$ is the X-features mean value.

Y_i is the Y-feature values in a sample.

$\overline{Y}$ is the Y-features mean value.

The Spearman correlation coefficient calculated by

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \tag{2}$$

Here ρ is the Spearman's rank correlation coefficient

d_i is the difference among the two ranks of each observation.

n is the total observations.

The correlation coefficient calculated by Kendall is

$$\tau = \frac{N_c - N_d}{\frac{n(n-1)}{2}} \tag{3}$$

Here τ is the Kendall rank correlation coefficient

N_c is the total of concordant.

N_d is the total of discordant.

Machine learning [8] is an area of AI and computer science that enables system models to access data and learn for themselves. Classification is a task that requires the use of machine learning techniques to figure out how to label problem domain objects with a class label. The difficulty of learning to locate records in a dataset that correspond to two or more target class labels is referred to as a classification challenge in Machine Learning. Multilayer perceptron (MLP) [9] is a feed-forward neural network. MLP contains input, hidden and output layers. Each layer contains a certain number of identical neurons. The input layer is the initial layer, and its neurons collect data from the input dataset features. The output layer is the final layer, and it contains one neuron for each value that the network produces. All of the layers in between input and output layers are called hidden layers because we do not know what these units should calculate ahead of time and have to figure it out while learning.

Optimization is the process of minimizing or maximizing any mathematical expression. By minimizing the function used to change the attributes of the neural networks, optimization methods are used to solve optimization problems. This study evaluates the SGD, LBFGS, and ADAM optimization methods with a multilayer perceptron classification algorithm for DDoS attack detection.

Adaptive Moment Estimation (ADAM) optimization method is combination of momentum and Root Mean Squared Propagation (RMSP). This method is convergence rapidly and very fast. It rectifies the vanishing learning rate and high variance. It is a default optimization method for MLP classification algorithm. Computationally it is more expensive.

SGD (Stochastic Gradient Descent) is a best way to fit linear classifiers to convex loss functions. SGD is used to apply large scale machine learning problems. The advantages of SGD are efficiency and ease of implementation. SGD requires a number of hyperparameters and iterations. It is affected by feature scaling.

LBFGS (Limited memory Broyden Fletcher Goldfarb Shanno algorithm) is an optimization algorithm of the quasi-Newton family methods using a limited amount of computer memory. It is more suitable for solving problems having large number of features. It does not require extensive hyperparameter tuning. Compare to SGD, it requires more memory and more iterations.

3 Results and Discussion

Experiments on a UDP flood DDoS attack detection are evaluated in this study. Removed missing value and infinity value records in pre-processing. By using a variance threshold filter-based feature selection method, removed constant and quasi-constant features from the dataset. UDP flood dataset contains 12 constant and 8 quasi-constant features. Now, independently apply the Pearson, Spearman, and Kendall correlation algorithms to the UDP dataset and find correlated features using a threshold value of ≥ 80. Pearson, Spearman, and Kendall correlation techniques identified 33, 43, and 41 correlated features in a UDP flood data set, respectively. Select uncorrelation feature subsets of correlation methods by removing correlation feature subsets of the feature set of UDP flood dataset. On a UDP flood dataset, Pearson, Spearman, and Kendall correlation techniques detected 26, 16, and 18 uncorrelated features, respectively.

This study evaluates Syn flood DDoS attack detection with three proposed uncorrelated feature subsets using accuracy, specificity, log loss, K-Fold cross validation and ROC-AUC score valuation metrics.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{4}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

$$F1\text{score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{8}$$

where TP is TRUE POSITIVE, TN is TRUE NEGATIVE, FP is FALSE POSITIVE, FN is FALSE NEGATIVE.

$$\text{Log}-\text{loss} = -\frac{1}{n}\sum_{i=1}^{n}[y_i \ln P_i + (1 - y_i)\ln(1 - P_i)] \tag{9}$$

Here, the number of observations is “n,” the predicted probability is “P,” and the actual value is “y.”

The K-Fold Cross (KFC) validation dataset is divided into K folds, among them one fold serving as a test and the rest folds serving as training. The model is then evaluated. This procedure is repeated until all folds have been chosen as a test.

The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) is a probability curve that displays the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold levels in order to determine the efficiency of model classification.

Table 1 shows the accuracy results of UDP-flood DDoS attack Detection with MLP classification algorithm with different optimization techniques using different uncorrelation feature subsets. MLP with LBFGS optimization produces best accuracy with Pearson uncorrelation feature subset. MLP with SGD optimization produces lowest accuracy with Kendall uncorrelation feature subset. Remaining all combinations produces the same accuracy values.

Table 2 shows the K-fold cross-validation accuracy results of the UDP-flood DDoS attack detection with MLP classification algorithm with different optimization techniques using different uncorrelation feature subsets. MLP with LBFGS optimization produces best K-fold cross-validation accuracy with Pearson uncorrelation feature subset. Among three MLP three optimization methods, ADM produces better K-fold cross-validation accuracy with all uncorrelated feature subsets. Among uncorrelated feature subsets, Pearson feature subset produces better K-fold cross-validation accuracy values with all MLP optimization techniques.

Table 3 shows the specificity results of the UDP-flood DDoS attack detection with MLP classification algorithm with different optimization techniques using different uncorrelation feature subsets. MLP with SGD optimization produces best specificity value with Pearson and Kendall uncorrelation feature subsets. MLP with SGD

Table 1 Overall model accuracy of the MLP classification algorithm on a UDP DDoS attacks dataset with different un-correlated feature subsets

Optimization techniques	Pearson	Spearman	Kendall
ADAM	99.92	99.92	99.92
SGD	99.92	99.92	99.90
LBFGS	99.94	99.92	99.92

Table 2 K-fold cross-validation accuracy scores of the MLP classification algorithm on a UDP attack dataset with different uncorrelated feature subsets

Optimization techniques	Pearson	Spearman	Kendall
ADAM	99.9345% (0.0060%)	99.9244% (0.0056%)	99.9267% (0.0062%)
SGD	99.9367% (0.0086%)	99.9216% (0.0082%)	99.9281% (0.0062%)
LBFGS	99.9338% (0.0074%)	99.9262% (0.0060%)	99.9299% (0.0050%)

Table 3 The specificity of the MLP classification algorithm on a UDP attacks dataset with different uncorrelated feature subsets

Optimization techniques	Pearson	Spearman	Kendall
ADAM	0.62	0.65	0.63
SGD	0.84	0.61	0.83
LBFGS	0.68	0.63	0.63

Table 4 The log–loss value of the MLP classification algorithm on a UDP attack dataset with different un-correlated feature subsets

Optimization techniques	Pearson	Spearman	Kendall
ADAM	0.026786241153743545	0.02678619288187981	0.028069372847961776
SGD	0.02887103124261505	0.028390196616939485	0.03480559516261052
LBFGS	0.021813955453226286	0.027427803287632373	0.026946619619249816

Table 5 ROC-AUC scores of the MLP classification algorithm on an UDP attack dataset with different uncorrelated feature subsets

Optimization techniques	Pearson	Spearman	Kendall
ADAM	0.9972597975194718	0.9997800829975707	0.9998124377631585
SGD	0.999789122780093	0.9997757957261204	0.9997877152830096
LBFGS	0.9998275315498715	0.9997817437263468	0.9997905302771762

optimization produces lowest specificity value with Spearmanl uncorrelation feature subsets.

Table 4 shows the log-loss values of the UDP-flood DDoS attack detection with MLP classification algorithm with different optimization techniques using different uncorrelation feature subsets. MLP with LBFGS optimization produces best log-loss value with Pearson uncorrelation feature subset. Among three uncorrelation feature subsets Pearson produces better log-loss values with all MLP three optimization methods. Among three uncorrelation feature subsets Kendall produces lowest log-loss values with all MLP three optimization methods.

Table 5 shows the ROC-AUC scores of the UDP-flood DDoS attack detection with MLP classification algorithm with different optimization techniques using different uncorrelation feature subsets. Kendall uncorrelation feature subset produces best ROC-AUC score value with all MLP with all optimization methods. Spearman and Kendall uncorrelation subsets produce the same ROC-AUC scores with all MLP optimization methods (Figs. 1, 2 and 3).

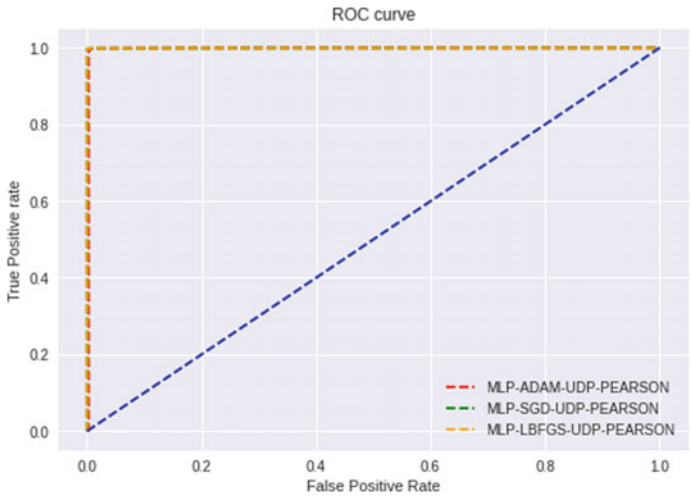


Fig. 1 ROC curves of the MLP classification algorithm on a UDP attack with Pearson uncorrelated feature subset

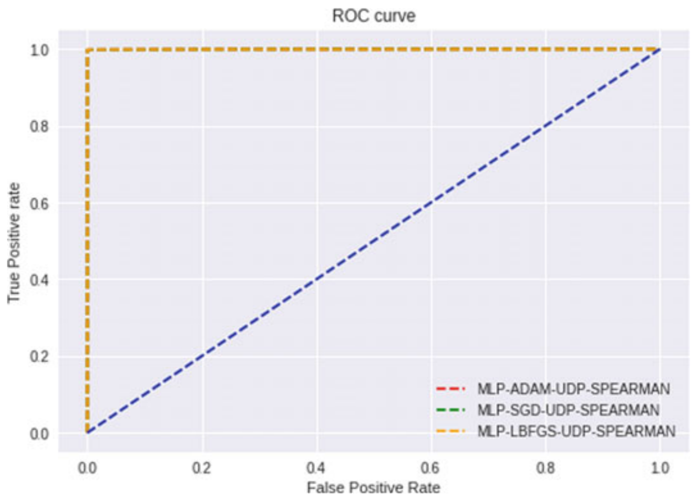


Fig. 2 ROC curves of the MLP classification algorithm on a UDP attack with Spearman uncorrelated feature subset

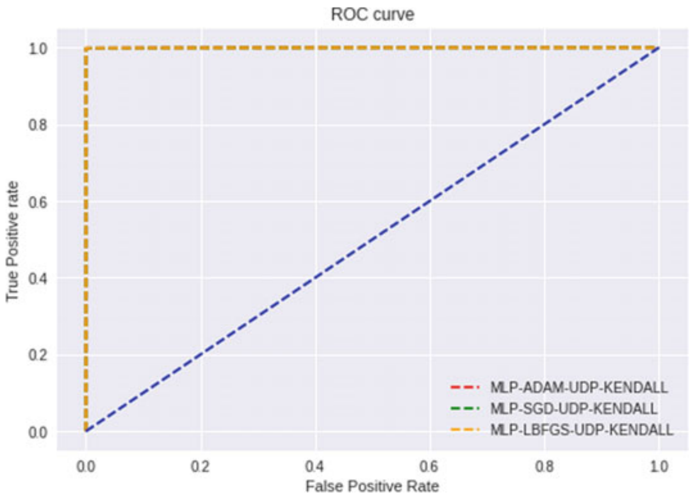


Fig. 3 ROC curves of the MLP classification algorithm on an UDP attack with Kendall uncorrelated feature subset

4 Conclusion

This research evaluates the Multilayer Perceptron classifier with the ADAM, SGD, and LBFGS optimization methods to classify UDP flood DDoS attacks and benign label classes. This work proposes using the Pearson, Spearman, and Kendall correlation algorithms to select a common uncorrelated feature subset for UDP flood DDoS attacks classification. MLP classifier all optimization methods gives better results on the UDP DDoS attack dataset with all uncorrelation feature subsets. MLP classifier gives best results with the SGD optimization method with all uncorrelation feature subsets on the UDP DDoS attack dataset. The extent of this research on UDP flood DDoS attack detection with MLP classification algorithm using features selected by KPCA [10] dimensionality reduction.

References

1. Dasari KB, Devarakonda N (2022) Detection of DDoS attacks using machine learning classification algorithms. International Journal of Computer Network and Information Security (IJCNIS) 14(6):89–97. https://doi.org/10.5815/ijcnis.2022.06.07
2. Dasari KB, Devarakonda N (2021) Detection of different DDoS attacks using machine learning classification algorithms. Ingénierie des Systèmes d'Information 26(5):461–468. https://doi.org/10.18280/isi.260505
3. Dasari KB, Devarakonda N (2022) Detection of TCP-based DDoS attacks with SVM classification with different kernel functions using common uncorrelated feature subsets. International Journal of Safety and Security Engineering 12(2):239–249. https://doi.org/10.18280/ijsse.120213

4. Al-Naymat G, Al-Kasassbeh M, Al-Harwari E (2018) Using machine learning methods for detecting network anomalies within SNMP-MIB dataset. International Journal of Wireless and Mobile Computing 15(1):67–76
5. Dasari KB, Devarakonda N (2022) SynFlood DDoS attack detection with SVM Kernels using uncorrelated feature subsets selected by Pearson, Spearman and Kendall Correlation Methods. In: 2022 second international conference on computer science, engineering and applications (ICCSEA), pp 1–6. https://doi.org/10.1109/ICCSEA54677.2022.9936114
6. Xiao P et al (2015) Detecting DDoS attacks against data center with correlation analysis. Computer Communications 67:66–74
7. Dasari KB, Devarakonda N (2022) TCP/UDP-based exploitation DDoS attack detection using AI classification algorithms with common uncorrelated feature subset selected by Pearson, Spearman and Kendall correlation methods. Revue d'Intelligence Artificielle 36(1):61–71. https://doi.org/10.18280/ria.360107
8. Tuan TA, Long HV, Son LH, Kumar R, Priyadarshini I, Son NTK (2020) Performance evaluation of Botnet DDoS attack detection using machine learning. Evolutionary Intelligence 13:283–294
9. Wang M, Lu Y, Qin J (2020) A dynamic MLP-based DDoS attack detection method using feature selection and feedback. Computers & Security 88:101645. ISSN 0167-4048. https://doi.org/10.1016/j.cose.2019.101645
10. Mekala S, Padmaja Rani B, Padmaja Rani B (2020). Article ID: IJARET_11_11_121 Kernel PCA based dimensionality reduction techniques for preprocessing of Telugu text documents for cluster analysis. International Journal of Advanced Research in Engineering and Technology 11(11):1337–1352. https://doi.org/10.34218/IJARET.11.11.2020.121

Intelligent Ensemble Algorithm for Feature Selection and Effective Prediction for Heart Disease Using SVM and KNN

R. Suresh and Nagaratna P. Hegde

Abstract Artificial intelligence capabilities are attracting the attention of many researchers and physicians who want to speed up the diagnosis process, aid in medical decisions, and deliver therapy sooner. Medical professionals could witness benefits in early intervention and patient outcomes for heart disease, cancer, and diabetic patients if they use AI. India needs its own data on heart disorders, which it can combine with cutting-edge technical methods to help detect those who are at risk of developing heart disease in the future. The goal of our research is to predict the existence of heart disease more accurately with a smaller no of attributes and including most prominent new attributes Troponin and B-type natriuretic peptide by employing an ensemble method based on Random Forest and a genetic algorithm to select the most important features. The accuracy obtained by the model is 99.36% by SVM and KNN algorithm outperformed other models.

Keywords Artificial Intelligence · Ensemble model · Feature selection

1 Introduction

Integrating Artificial Intelligence (AI) and Machine Learning (ML) to diagnose cardiac illnesses early was one of the major topics discussed at the Cardio logical Society of India's 73rd annual conference [2]. The CSI collaborated with scientists and engineers to construct "Deep Neural Networks," with the goal of effectively identifying cardiac risks in advance by 2030. It is eager to collaborate with the government in order to achieve the goal quickly. Heart attacks strike rate of South Asians is 10 years earlier than other ethnic groups [1]. The percentage of persons

R. Suresh (✉)
Department of CSD, Ace Engineering College, Hyderabad, India
e-mail: rella.suresh@gmail.com

N. P. Hegde
Department of CSE, Vasavi College of Engineering, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_53

under 40 or 45 years old who had heart attacks was little 10–15 years ago, about 5–7%, but some of the new data is worrying. When doctors treat patients who are at risk for heart disease, race has a role in the treatment they receive [3]. This isn't always about prejudice. Black patients, according to studies, have a higher risk of having heart problems.

In patients with or at high risk of cardiovascular disease, risk stratification is routinely used to customize care aimed at preventing heart attacks, strokes, and sudden cardiac death. Traditional calculators only take into account a small number of clinical data, such as age, gender, smoking status, blood pressure, and cholesterol [4]. The work differentiated the performance of machine learning with current scores in predicting cause death in individuals with suspected or established coronary artery disease using Tropt and Bnp clinical data [8]. Feeling well despite the absence of sickness signs is a self-determination based on a lack of actual information of our body' state. We only went to the doctor when we felt sick. Now, doctors can tell us whether we are disease-free or not based on machine learning techniques.

2 Literature Review

Because the nature of the heart is complex, it must be handled with care, or the individual will die. Various methods, such as KNN, DT, GA and NB, are used to classify the severity of cardiac disorders [9]. Mohan et al. outline how two independent approaches can be combined to form a single hybrid strategy with an accuracy of 88.4%, it is a elevated approach than all other models [5].

For the prediction of heart disease, author Ashraf et al. used ML algorithms as well as ensemble models such as Bayes Network, J48, KNN, NB, RT, and RF. J48 was the most accurate, with a score of 70.77% [6].

Polaraju suggested Prediction of Heart Disease Using MRM, demonstrating that MLR is suitable for forecasting the possibility of heart disease. The work based on the training data which includes 3000 records of patients with the 13 different features described earlier. The data set is separated into two sections, with 70% utilized for training and 30% for testing [7].

3 Methodology and Modeling

To protect against over fitting, use cross-validation prior to modeling. The data set was divided into two sections, with an 80–20% split [10]. The training set is taken 80% of the data used to train the model, while the testing set is taken 20% of dataset utilized to validate it.

Random forest is a supervised ML technique that is widely used. Before we go into the random forest concept. Let's start by looking at the components of random forest decision trees [11]. A decision tree is a kind of ML algorithm which splits dataset into categories. Division starts by performing binary split which carried out till no further splits are able to be done [12]. Different branches of varying lengths are developed. DT is to precipitate the training dataset into a compactable feasible tree. The logical rule for a grouping of incidence is favored over other clarifications is rationale for minimizing tree size.

There are numerous advantages to using decision trees. It features simple visualization, for example, that is easy for people to grasp and interpret. Building decision trees, on the other hand, has significant drawbacks. This approach is overly greedy it necessitates the capability of an algorithm to discover pre-eminent choice at each node [13]. It satisfies each step's optimal option while also sacrificing the global optimum. Furthermore, especially when a tree is exceptionally deep, decision trees are prone to over fitting.

Random forest, on the other hand, addresses this issue by training on diverse samples of data sets using a random subset of characteristics. It is build with large number of independent DT which exists with each other to form an ensemble [14]. Each tree in the RF eliminates a class prediction, and the class with the highest votes is considered for classification model. RF is an ensemble learning method that addresses the disadvantages of decision trees' tendency to over fit their training set.

The data set of 10,000 records with 16 attributes used for consistency. Categorical attributes are used by classification models. The attributes from disease dataset are reduced to sixteen using Ensemble Genetic algorithm. The selected attributes data set is given to as input to classification algorithms. K-fold cross validation technique is employed to test classification models accuracies. The target attribute is given with value "0" represent negative to cardiac disease and value "1" to represent positive of cardiac disease.

In predicting the dependent variable "Target," two variables play a crucial influence. That's "tropt" (troponin level) and "BNP" (brain natriuretic peptide) (Brain natriuretic peptide). To study and monitor heart failure, employed Tropt and Bnp with other important measurements. Both high tropt and high Bnp readings, according to the study, can cause a heart attack (Fig. 1).

The 16 attributes which were selected from a total of 73 attribute attributes of heart disease patients dataset using Ensemble genetic model in python programming language (Fig. 2).

From the above bar chart out of total 1040 patient's records found 626 people suffering from heart disease and 414 not suffering from any heart disease (Fig. 3; Tables 1, 2, 3 and 4).

The classification algorithms considered for cardiac disease prediction by use of ensemble model for feature selection obtained an accuracy of 97.68 for Naive Bayes, 99.36 by SVM algorithm and 99.36 by KNN algorithm.

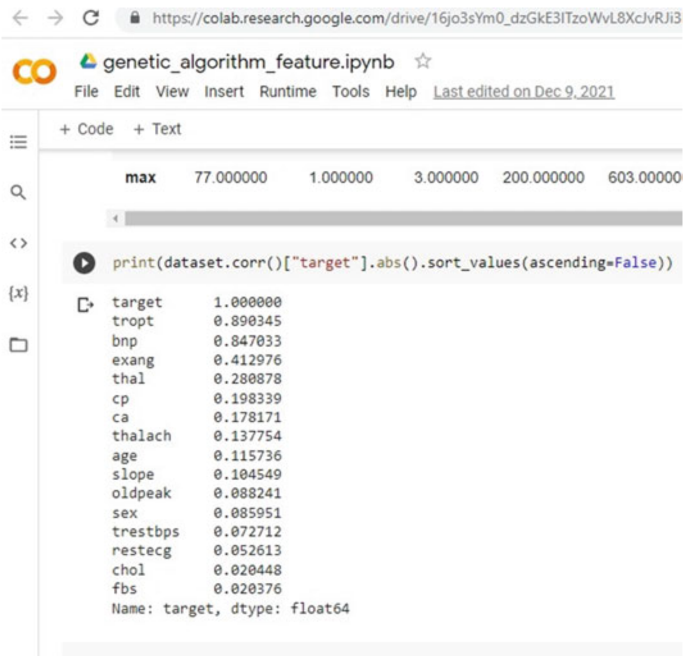


Fig. 1 Attributes selected based on feature selection using Ensemble model for feature selection

Fig. 2 Heart disease prediction using python from 1043 records

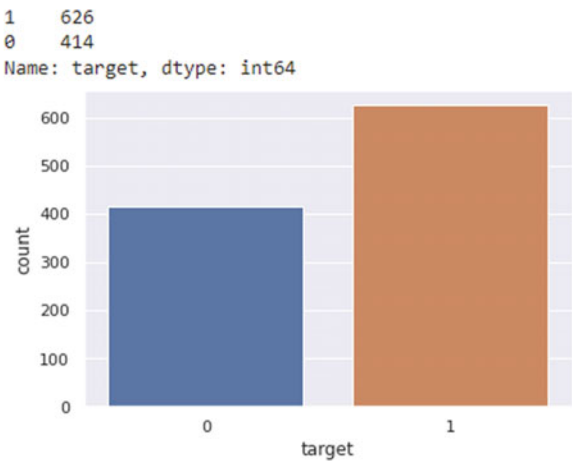


Fig. 3 Accuracies obtained by various prediction algorithms

Table 1 Classification report obtained by Naïve Bayes algorithm with various performance parameters

Algorithm used	Class	Precision	Recall	F-score	AP	Accuracy
NB	0	0.89	0.89	0.89	0.88	90.16
	1	0.91	0.91	0.91		
Ensemble GA + NB	0	0.98	0.97	0.97	0.97	97.68
	1	0.96	0.99	0.98		

Table 2 Classification report obtained by SVM algorithm with various performance parameters

Algorithm used	Class	Precision	Recall	F-score	AP	Accuracy
SVM	0	0.88	0.81	0.85	0.83	86.89
	1	0.86	0.91	0.89		
Ensemble GA + SV M	0	0.97	0.96	0.97	0.99	99.36
	1	0.98	0.99	0.99		

Table 3 Classification report obtained by KNN algorithm with various performance parameters

Algorithm used	Class	Precision	Recall	F-score	AP	Accuracy
KNN	0	0.58	0.56	0.57	0.62	62.30
	1	0.66	0.68	0.67		
Ensemble GA + KN N	0	0.99	0.98	1.00	1.0	99.36
	1	0.99	0.98	1.00		

Table 4 Accuracy obtained by various algorithms

Serial No.	Algorithm	Accuracy (%)
1	NB	90.16
2	SVM	86.89
3	KNN	62.3
4	Ensemble Genetic + NB	97.68
5	Ensemble Genetic + SVM	99.36
6	Ensemble Genetic + KNN	99.36

4 Conclusion

Our goal is to use a smaller set of variables to predict the existence of heart disease more accurately. Initially, 13 characteristics were used to predict heart disease. In our research, a genetic algorithm is utilized to identify the characteristics that contribute the most to the diagnosis of cardiac problems, hence reducing the amount of tests that a patient must do. Using genetic search, 73 attributes are reduced to 16 attributes. When we look at the model from the standpoint of accuracy, we can see that the accuracies of all three models for the testing data set are above 80%, which is quite good. Following that, three classifiers such as Naive Bayes, SVM, and KNN are utilized to forecast the Cardiac patients with the highest accuracy once the number of characteristics is reduced. Furthermore, the results show that after including feature subset selection and a small delay in construction time, the SVM and KNN data mining technique beats the other data mining strategies. With the same model development time, Naive Bayes performed better after attribute pruning. Unpredictable and missing values were handled before the classification models were built. The goal of this project is to build on previous work by using fuzzy learning models to assess the severity of cardiac illness.

References

1. Al-Makhadmeh Z, Tolba A (2019) Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: a classification approach. Measurement 147:106815
2. Ali F, Islam SMR, Kwak D, Khan P, Ullah N, Yoo SJ, Kwak KS (2017) Type-2 fuzzy ontology-aided recommendation systems for IoT-based healthcare. Comput Commun 0–1
3. Ali F, Kim EK, Kim Y-G (2015) Type-2 fuzzy ontology-based semantic knowledge for collision avoidance of autonomous underwater vehicles. Inf Sci (Ny) 295:441–464
4. Ali F, Kim EK, Kim YG (2015) Type-2 fuzzy ontology-based opinion mining and information extraction: a proposal to automate the hotel reservation system. Appl Intell 42:481–500
5. Ali F, Kwak D, Khan P, El-Sappagh S, Ali A, Ullah S, Kim KH, Kwak K-S (2019) Transportation sentiment analysis using word embedding and ontology-based topic modeling. Knowledge-Based Syst

6. Ali F, Kwak D, Khan P, Islam SMR, Kim KH, Kwak KS (2017) Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling. Transp Res Part C Emerg Technol 77:33–48
7. Ali F, Kwak K-S, Kim Y-G (2016) Opinion mining based on fuzzy domain ontology and support vector machine: a proposal to automate online review classification. Appl Soft Comput 47:235–250
8. Ayata D, Yaslan Y, Kamasak ME (2018) Emotion based music recommendation system using wearable physiological sensors. IEEE Trans Consum Electron 64:196–203
9. Bernal EA, Yang X, Li Q, Kumar J, Madhvanath S, Ramesh P, Bala R (2018) Deep temporal multimodal fusion for medical procedure monitoring using wearable sensors. IEEE Trans Multimed 20:107–118
10. Bhuvaneswari ANG (2014) An intelligent approach based on principal component analysis and adaptive neuro fuzzy inference system for predicting the risk of cardiovascular diseases. In: 2013 5th international conference on advance computing, ICoAC 2013, pp 241–245
11. Din S, Paul A (2019) Smart health monitoring and management system: toward autonomous wearable sensing for Internet of Things using big data analytics. Futur Gener Comput Syst 91:611–619
12. Fazlic LB, Hallawa A, Schmeink A, Peine A, Martin L, Dartmann G (2019) A novel NLP-fuzzy system prototype for information extraction from medical guidelines. In: Proceedings of 2019 42nd international Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2019), pp 1025–1030
13. Hao Y, Usama M, Yang J, Hossain MS, Ghoneim A (2019) Recurrent convolutional neural network based multimodal disease risk prediction. Futur Gener Comput Syst 92:76–83
14. Jabeen F, Maqsood M, Ghazanfar MA, Aadil F, Khan S, Khan MF, Mehmood I (2019) An IoT based efficient hybrid recommender system for cardiovascular disease. Peer-to-Peer Netw Appl 12:1263–1276

Efficient Brain Tumour Classification Using Parameter Optimized CNN with Dingo Optimizer Concept

R. Aishwarya, Ganesan Sumathi, and T. K. S. RathisBabu

Abstract Brain tumour is a serious issue caused in the brain, where a group of unwanted tissues grows in unrestrained manner. The image-processing field has gained outstanding growth in the region of medical applications with the discovery of various approaches in baseline methods. Convolutional Neural Networks (CNN) is applied in the process of brain tumour identification and categorization due to their significant performance in image processing. The main limitation of existing approaches is that it has a large count of constraints and produced elevated implementation period and requires high specification for system implementation. Hence, this paper aims to design the efficient brain tumour classification with the help of various technologies. At first, the data is collected from online resources. Next, it is pre-processed by the 'median filtering and the Contrast Limited Adaptive Histogram Enhancement (CLAHE) techniques'. Then, the pre-processed images are inputted to the final classification phase, where the final output is classified by the parameter optimized CNN by Dingo Optimizer (DOX). The results on the dataset have revealed the best performance in comparison with several traditional methods.

Keywords Brain Tumour Classification · Convolutional Neural Network · Contrast Limited Adaptive Histogram Enhancement · Dingo Optimizer · Parameter optimized CNN

R. Aishwarya · G. Sumathi (✉)
Faculty of Engineering and Technology, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Cuddalore, Tamil Nadu, India
e-mail: sumi.ganesan@yahoo.com

T. K. S. RathisBabu
Department of Computer Science and Engineering, Sridevi Women's Engineering College, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_54

1 Introduction

The brain is a vital organ in human body with compound structures [1]. The brain is surrounded with the skull region that makes it difficult to detect activities of brain. Abnormality of brain would lead in affecting structure of brain and cause brain cancer. The report of World Health Organization (WHO) has specified that 9.6 million peoples overall died of cancer in 2018 [2]. For analysing the brain tumour, 'Magnetic resonance imaging (MRI)' is usually used to spot the brain tumours with high-quality images. The MRI provides the best visualization of both contrasts and spatial resolution [3].

Brain is a complicated organ, which play an essential part in our daily behaviours and the uncontrollable, unequal tissue growth in the brain is named as brain tumour that is a dreadful disease [4]. The most widely detected tumour types are Pituitary, Meningioma, and Glioma [5]. The Glioma tumours are high-grade tumours usually present in cerebral hemisphere named as attocycles, meningioma is low-grade tumours in skull, Tumours in pituitary gland are called pituitary tumour [6].

The recent researches were undergone to implement machine learning to detect the Region of Interest (ROI) in data in an efficient manner than other traditional methods [7]. Baseline algorithms have the capability to identify different ranges of data depiction for the detection and prediction task. CNN is a type of deep neural networks that is utilized effectively on image-processing technique that involves object detection, segmentation, and classification of image [8]. The CNN is concentrated mostly in object detection and classification. Numerous researches have executed in the region of brain tumour identification, segmentation, categorization and various positive results have attained based on this techniques. The other techniques like 'Statistical Region Merging, K-means clustering and Fuzzy Clustering' are new approaches in predictive analysis for the identification and categorization of brain tumour, and these approaches are performed better than other conventional approaches [9]. The adoption of deep learning methods is often allocated with high computational cost and reduction in accuracy [10]. Moreover, it has less prediction certainty that keeps the life of human in loop, which is dangerous for medical field. The main contribution of brain tumour classification method is explained below. To design an efficient brain tumour classification by inputting MRI brain images with CNN for classification and optimization-based algorithm for parameter tuning. Categorize the brain images with CNN and optimization of hidden neurons and count of epochs using DOX algorithm and by increasing the accuracy and precision of brain tumour classification. Demonstrate the performance of the introduced method over various existing methods by analysing numerous paradigm measures.

2 Literature Survey

2.1 Related Works

In 2020, Ge et al. [11] have investigated the issues in the classification of brain tumour with MRI from various scanning procedures like 'T2 weighted, T1 weighted', FLAIR and contrast improved with T1-weighted images. The currently available glioma dataset was different in dimension and mostly associated with unfinished MRIs in dissimilar modalities. To overcome the issues related to insufficient dataset and unfinished modality of brain images, a new Generative Adversarial Network (GAN) method was proposed to expand the brain tumour dataset. The GAN was capable to produce artificial MRIs in different procedures. The post-processing method was introduced to detect the glucose level of patients. The attributes of glioma were extracted with two phase GAN-improved MRIs. The effectiveness of the suggested scheme was estimated and the experimental were performed on brain tumour dataset for glioma classification. The simulation results have achieved good performance than other existing approaches.

In 2021, Habib et al. [12] have planned for the segmentation of brain tumour on MRI images with threshold segmentation and then brain tumour images were classified, and then the features were extracted with various classifiers. The suggested methodology has involved four phases like 'image acquisition, image pre-processing, image segmentation, and feature extraction'. The brain tumours datasets were classified accurately with various classifiers. The conclusions have specified that the recommended framework has attained high accuracy in brain tumour detection. In 2021, Khairandish et al. [13] have categorized the MRI brain images using CNN in standard dataset to categorize Malignant and Benign tumours. The deep learning models have achieved high performance in image classification in recent years. The CNN approach was implemented to extract the features from images and has attained high accuracy regarding classification and for the threshold-based segmentation regarding classification.

In 2021, Kesav and Jibukumar [14] have suggested a novel framework for brain tumour object recognition and categorization via RNN approach, which was analysed on two standard datasets from Kaggle and figshare. The proposed work has planned to reduce the training period of RNN with less complex scheme and implemented a novel framework for brain detection. The CNN was adopted to categorize the healthy and unhealthy tumours in MRI images. Then, the attributes were extracted with RCNN to identify the brain tumour regions with MRI images and further, the region of tumour was enclosed with bouncing boxes. Therefore, this approach has expanded to notice various types of tumours. Hence, the introduced model has attained low training time than other existing techniques. In 2020, Afshar et al. [15] have suggested 'Bayesian CapsNet' scheme that was utilized to predict the brain tumour and also measure the improbability in prediction. The obtained results have shown that the removing the improbability in forecast has enhanced the precision of the recommended method. The removal of undefined prediction was a suitable approach for enhancing the network interpretability.

2.2 Problem Statement

The CNNs, which were the most image-based applications, are susceptible to lose essential spatial details among image occurrences. Because of the lethal nature of this malignancy and the repercussions of tumour misclassification, brain tumour classification is of fundamental relevance in the field of medical imaging difficulties. GAN [11] is robust and efficient and the performance of classification is better. But multiple datasets are not applied. Hybrid algorithm [12] utilizes the datasets in an appropriate manner and the accuracy is very high. Still, the exact location and size of the tumour is not considered. CNN-SVM [13] offers much efficient and enhancement approaches for the classification and high classification output rate is attained. Yet, various optimization algorithms are not considered. RCNN [14] supports in the real time processing and can be incorporated in handheld tools having lesser computational facilities. But it has the drawback of object detection. BayesCap [15] increases the trust into the decisions and the accuracy is enhanced. Still, the aleatoric uncertainty is not included. Hence, it is needed to include new deep structured learning approaches for efficient brain tumour classification.

2.3 Proposed Model- Brain Tumour Classification Enabled by Proposed Deep Learning

Brain is the majority compound organ with billions of neuron cells. The brain tumours are formed when the uncontrolled cells divide into abnormal cell groups. The brain tumours are categorized into high grade and low grade tumours. The MRI method is usually used to detect the tumour cells in brain and it has a great capability to analyse the brain structures. However, the high grade cancers are difficult to detect because it is formed under the skull regions. Various learning methods were adopted to recognize the image accurately. CNN is used for detecting and classifying the brain tumours automatically. Therefore, optimized CNN is adopted in this proposed method for accurate classification of brain tumours. The structural representation of brain tumour classification is represented in Fig. 1.

The proposed method covers three stages like, '(a) data collection, (b) pre-processing, and (c) classification'. The dataset is selected from model dataset. At first the input images are subjected to pre-processing phase, which is employed with median filtering and CLAHE. Later, the pre-processed image is given to classification task using CNN. In the classification phase, the parameters such as hidden

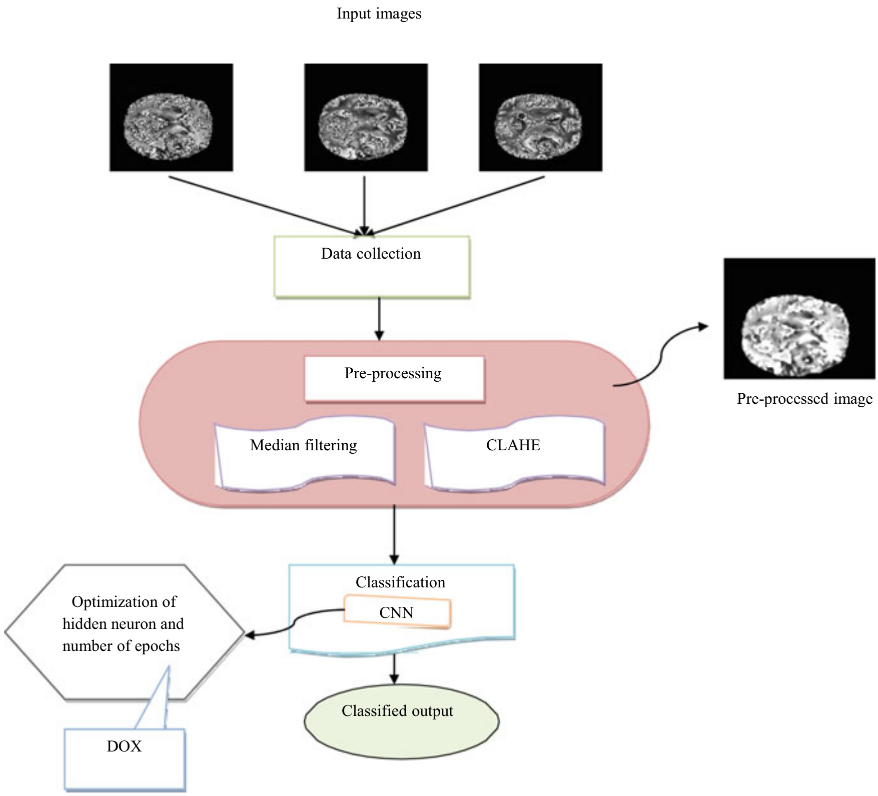


Fig. 1 Architectural representation for brain tumour classification

neuron and epoch count are optimized using DOX with the function of increasing the precision and accuracy of the brain tumour classification model.

2.4 Brain Tumour Classification Dataset

The dataset is selected from 'https://www.kaggle.com/mateuszbuda/lgg-mri-segmentation'. The brain MRI segmentation dataset compiles of MRI brain images with FLAIR and segmentation masks. The input images are gathered from 'The Cancer Imaging Archive (TCIA)', which is related to 110 adults with lower and higher grade glioma. The composed data are referred as Bi_t, where $t = 1,2,3,\ldots,T$ and the overall number of collected data is referred as T. The sample image of brain tumour is illustrated in Fig. 2.

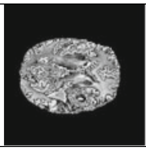
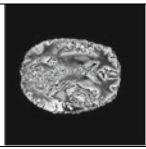
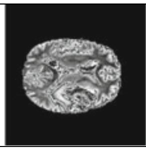
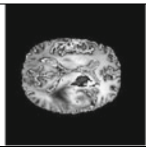
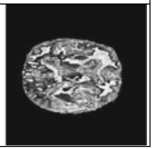
Descriptions	Image 1	Image 2	Image 3	Image 4	Image 5
Sample images					

Fig. 2 Sample images of brain tumour classification

3 Pre-processing and Classification of Brain Tumour Using Parameter Optimized CNN

3.1 MRI Image Pre-processing

The pre-processing is a crucial step in any processing of image models to enhance the contrast of the images and for the illumination changes. It is essential to progress the quality of image and amplify the noises present in the image. The pre-processing of MRI images is done with median filtering and CLAHE.

Median filtering**:** It is applied to the images to neglect the noise and to restrain the huge or small frequencies to further enhancement or to detect the edges of the images clearly. A non-linear filter is used in this filtering process. The main purpose of the median filter is to eliminate the noisy pixels by the median values of the surrounding pixels and removes unwanted things present in the in the segmented image. Some isolated pixels if image contains very low and high intensity values, this impulse noise is removed by median filtering. The median filtered images are referred as Bi^{median}.

CLAHE: The median filtered image Bi^{median} is inputted to CLAHE. The contrast enhancement is the effective technique utilized to amplify the noises troubles and to boost the contrast in the image. CLAHE is widely used in contrast enhancement and it is more effective in medical image processing. This technique works well than other regular histogram equalization, the value of the filter which is given on the neighboring pixel is computed and it does not count the global filter value. The final preprocessed images are derived as Bi^{pre}.

3.2 CNN-Based Classification

The brain tumour images are classified using CNN. It is a deep learning architecture used to recognize the special designs in 2D images. It has many layers like "convolutional layer, pooling layer, sample layer, input layer and output layer". CNN model consists of pooling region, fully connected region and convolutional region. The convolution region is dispersed with pooling region to decrease the consumption period and to construct enhanced spatial and configuration variance. Convolutional region has a rectangular rid with multiple neurons. The convolutional layer uses different segmented images as input. There is a pooling layer after convolutional layer, which is illustrated from the prior convolutional layer.

Convolutional layer: The convolutional layers evaluate the input images and act as a feature extractor. The neurons in convolutional layer are determined as feature maps. Both weights and inputs are combined together to initialize the features map and passed through the activation functions. The feature maps presented in the same region has distinct weights, thus from each regions, various features are extracted.

Pooling layer: It is also named as down sampling region. The input distortions and spatial invariance are obtained in extraction process. All input average values presented in a small image are moved to the next layer.

Fully connected region: It prefers the n dimensional input data, the term n establish the count of categories gathered for the program. The class possibility is represented with n dimensional matrix in all ranges. The precise probability from different classes is attained by fully connected region.

3.3 DOX-CNN-Based Classification

The proposed DOX-CNN-based classification is performed to categorize the MRI images with CNN and to optimize the parameters using DOX. The suggested method utilizes the DOX owing to its ability to solve all types of optimization issues and it selects best options for real life issues. The effectiveness of procedures is enhanced and has high convergence speed. DOX is inspired by the dingoes hunting activity. Usually, the dingoes hunt in cluster grouping. The hunting method is classified into three stages like 'approaching, chasing, encircling and attacking'. The major phases of DOX are exploitation and exploration phase. The optimal solutions are selected using the exploitation stage.

The location of the preys is easily detected by dingo. The alpha encircles the prey after finding its site. The prey is targeted by the preeminent agent and the surrounding activity of dingoes is derived here.

$$\vec{k}_r(u+1) = \beta_1 \sum_{0=1}^{n} \frac{\left[\varphi_o(u) - \vec{k}_r(u)\right]}{n} - \vec{k}_*(u) \tag{1}$$

Here, the variable $\vec{k}_r(u+1)$ represents the novel location of the search agent, n denotes the random integer, the sub set of candidate solution is indicated a $\varphi_{\text{o}}(u)$, the population generated is taken as U, the current candidate solution is termed as $\vec{k}_r(u)$, the best candidate solution from the prior iteration is denoted as $\vec{k}_*(u)$, the uniformly produced random number is represented as β_1 at the time interval $-$ 2 to 2.

The dingo usually chases the smaller victim and it follows the victim until it is trapped alone. The following activity of the dingo is expressed in Eq. (2).

$$\vec{k}_r(u+1) = \vec{k}*(u) + \beta_1 * x^{\beta_2} * \left(\vec{k}_e(u) - \vec{k}_r(u)\right) \tag{2}$$

The variables $\vec{k}_r(u+1)$ denotes the dingoes movement, $\vec{k}_r(u)$ explains the present iteration, the best candidate solution from the prior iteration is denoted as $\vec{k}_*(u)$, the arbitrary number β_1 lies in the gap $-$ 1 to 1 and the chosen candidate solution is referred as $\vec{k}_e(u)$.

The activity of scavenger is the tradition of dingo eating the victim, the forager activity of dingo is expressed in Eq. (3).

$$\vec{k}_r(u+1) = \frac{1}{2}\left[y^{\beta_2} * \vec{k}_e(u) - (-1)^{\sigma} * \vec{k}_r(u)\right] \tag{3}$$

The dingo's survival range is referred in Eq. (4).

$$\text{survival}(r) = \frac{ft_{\max} - ft(r)}{ft_{\max} - ft_{\min}} \tag{4}$$

The variables $ft_{\max}$ and $ft_{\min}$ are indicated as the worst and optimal fitness values in the present iteration. The variable $ft(r)$ indicates the present robustness value of rth search agent. The standardized fitness value is available in the survival matrix at bounding limit [0, 1]. Mention about Eq. 5.

$$\vec{k}_r(u) = \vec{k}*(u) + \frac{1}{2}\left[\vec{k}_{e_1}(u) - (-1)^{\sigma} * \vec{k}_{e_2}(u)\right] \tag{5}$$

The survival range of the candidate solution is represented as $\vec{k}_r(u)$, the terms e_2 and e_1 is the generated random numbers, respectively, $\vec{k}*(u)$ is indicated as the best candidate solution from the former iteration and the term σ is the binary number. The pseudo code for DOX is shown in Algorithm 1.

Algorithm 1: Introduced DOX
Population the initialization
Create the optimal location based on fitness solution
Estimate the fitness for all resolution
Conclude the constraints
Upgrade the latest candidate solution by Eq. (1)
Upgrade the chasing activity with Eq. (2)
Upgrade the activity of scavenger with Eq. (3)
Update the survival range with Eq. (4)
Find the preeminent solution

(continued)

(continued)

Algorithm 1: Introduced DOX
end

The pre-processed MRI images are classified with CNN and the parameters like hidden neuron number and number of epochs is tuned with DOX algorithm with the principle of enhancing the precision and accuracy of the categorization. The main objective of the suggested method is derived in Eq. (6).

$$ft = \underset{\{Hn,Ne\}}{\arg\min}\left(\frac{1}{\text{acu} + \text{per}}\right) \tag{6}$$

In above formula, the term *Hn* is denoted as hidden neuron count, the count of epochs is termed as *Ne*. The hidden neuron is taken among [5–255] and the epochs count ranges from [0.01–0.99]. "Accuracy referred as to how closely the measured value of a quantity corresponds to its true value".

$$\text{acu} = \frac{ab + an}{ab + an + ac + ad} \tag{7}$$

The variables, *ab*, *an*, *ac*, *ad* are referred as "true positive, true negative, false positive and false negative". "Precision is the positive predictive value or the fraction of the positive predictions that are actually positive".

$$\text{pre} = \frac{ab}{ab + ad} \tag{8}$$

Thus, the proposed DOX-CNN-based classification phase increases the accuracy and precision of the suggested method.

4 Results and Discussions

4.1 Experimental Analysis

The introduced brain tumour categorization was executed in Python and the analyses of the experiments were established. The performance of the projected method was evaluated with various traditional method regarding various metrics such as "accuracy, sensitivity, specificity, precision, Net Present Value (NPV), F1 Score, False-positive rate (FPR), false-negative rate (FNR), and False Discovery Rate (FDR)". The suggested approach is evaluated with other algorithms like PSO [16], DHOA [17], SSA [18], EHO [19] and further, evaluated with LSTM [20], SVM [21], DNN [5], CNN [22].

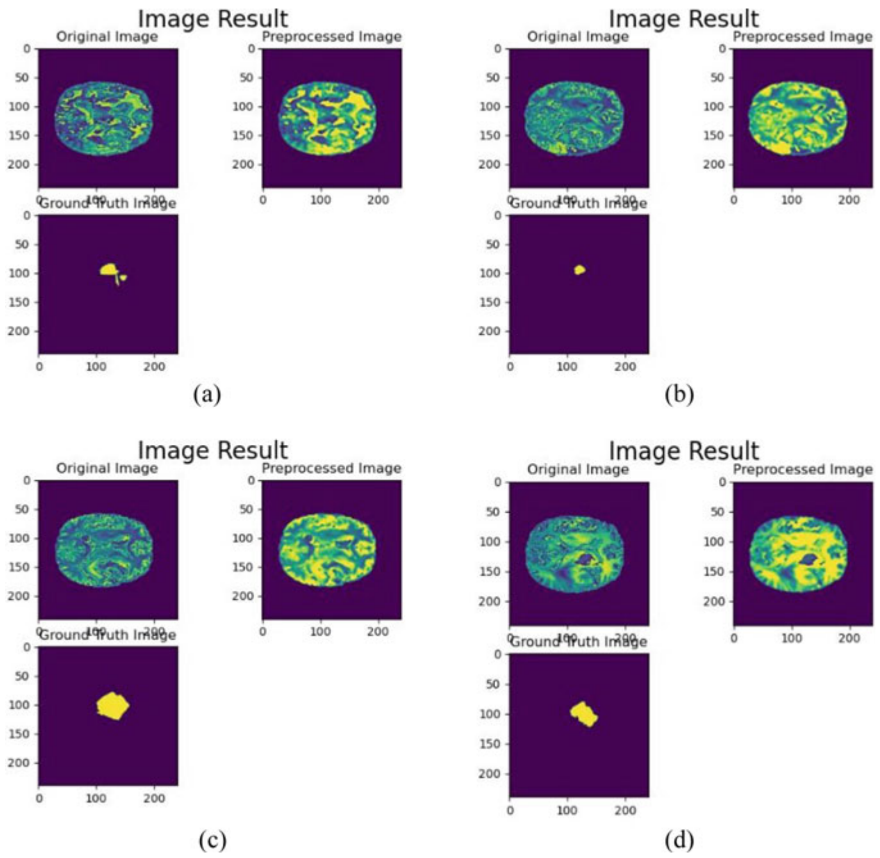


Fig. 3 Image results for introduced brain tumour classification

4.2 Image Results

The image results for the introduced brain tumour classification is given in Fig. 3.

4.3 Performance Measures

The performance measures of the introduced brain tumour classification are taken from "https://en.wikipedia.org/wiki/Sensitivity_and_specificity: Access Date: 2022-03-21".

4.4 Analysis of Performance on Algorithms

The analysis of modified brain tumour classification is assessed with algorithms by differing the learning rate is shown in Fig. 4. The accuracy of the modified DOX-CNN is 16% higher than PSO-CNN, 2% higher than DHOA-CNN, 3% higher than SSA-CNN, 4% higher than EHO-CNN at learning percentage 40%. While comparing the evaluated results, the precision of implemented DOX-CNN is 3% improved than PSO-CNN, 5% improved than DHOA-CNN, 4% improved than SSA-CNN, 6% improved than EHO-CNN at learning rate 50%. Thus, the proposed DOX-CNN is attaining best results than other traditional approaches.

4.5 Performance on Traditional Learning Models

The analysis measures of introduced brain tumour classification is executed with prior existing learning methods by differentiating learning rate as shown in Fig. 5. The accuracy of the implemented DOX-CNN is 44% improved then LSTM, 29% improved than SVM, 19% improved than DNN, 6% improved than CNN at the given learning rate 30%. Moreover, at learning rate 40%, the specificity of the suggested DOX-CNN is 48% higher than LSTM, 31% higher than SVM, 15% higher than CNN and 12% higher than DNN. Thus, the suggested method has gained high classification accuracy.

4.6 Performance on Algorithms

The analysis of introduced brain tumour classification is assessed with different algorithms is given in Table 1. The accuracy of the introduced DOX-CNN is 15% progressed than PSO-CNN, 12% progressed than DHOA-CNN, 13% progressed than SSA-CNN and 14% progressed than EHO-CNN. The precision of implemented DOX-CNN is 13% progressed than PSO-CNN, 15% progressed than DHOA-CNN, 14% progressed than SSA-CNN and 16% progressed than EHO-CNN. Thus, the proposed DOX-CNN has gained best results than other algorithms.

4.7 Performances on Traditional Learning Models

The analysis metrics of introduced brain tumour classification is evaluated with other existing learning methods by differentiating learning rate as shown in Table 2. The accuracy of the suggested DOX-CNN is 4% higher then LSTM, 9% higher than SVM, 6% higher than DNN, 8% higher than CNN. The specificity of the suggested

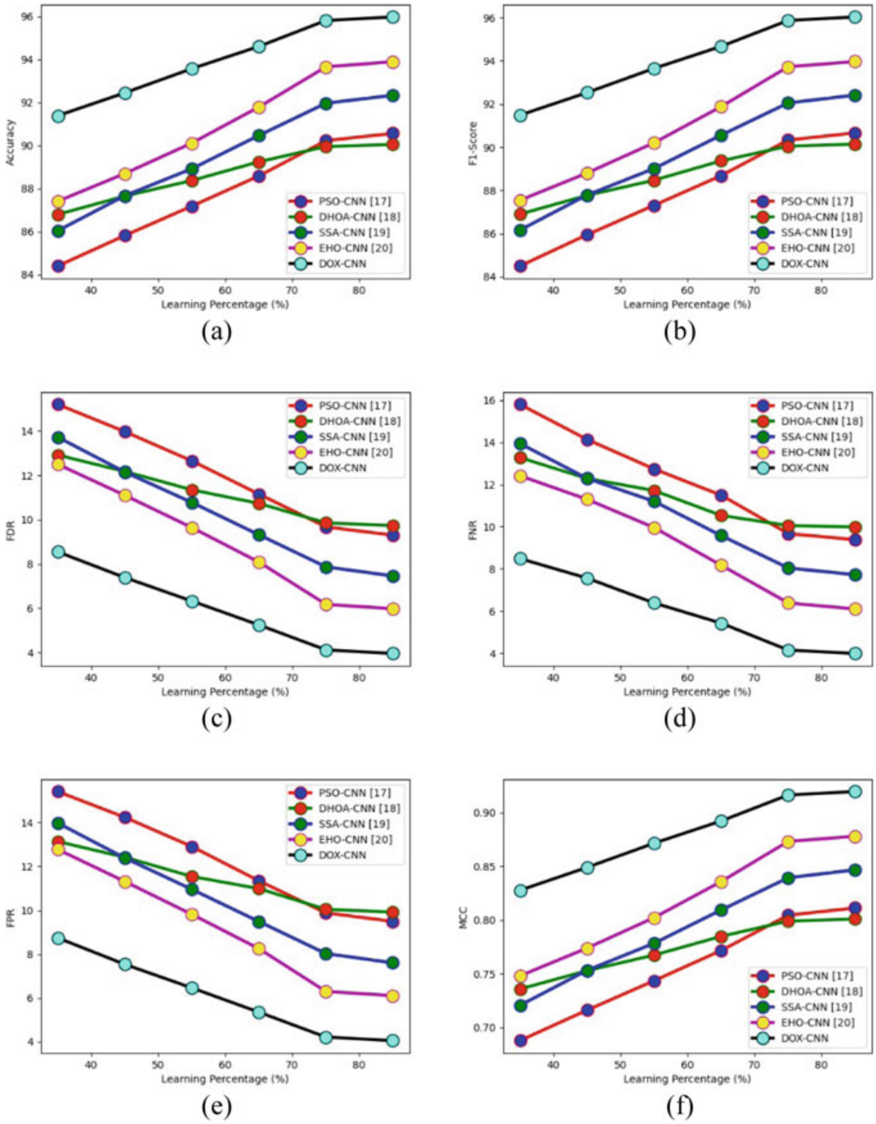


Fig. 4 Analysis of recommended brain tumour classification on traditional learning models at different learning rate concerning "**a** Accuracy, **b** F11-score, **c** FDR, **d** FNR, **e** FPR, **f** MCC, **g** NPV, **h** Precision, **i** Sensitivity, **j** Specificity"

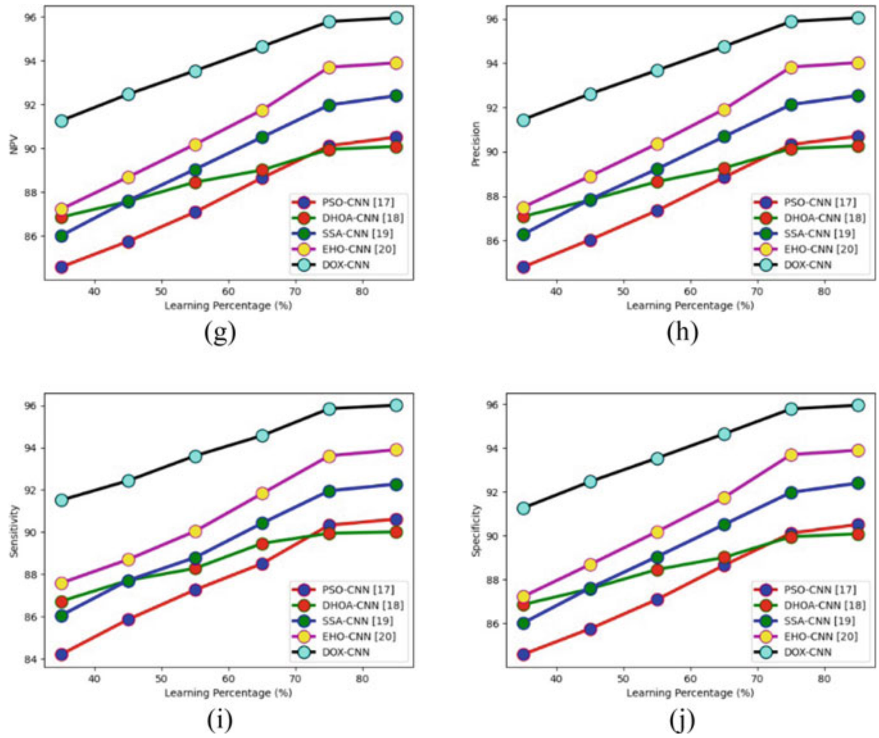


Fig. 4 (continued)

DOX-CNN is 8% enhanced than LSTM, 3% enhanced than SVM, 11% enhanced than DNN and 15% enhanced than CNN. Thus, the suggested method has gained high classification accuracy.

5 Conclusion

The proposed brain tumour classification was implemented with a CNN enabled with optimization algorithm. The collected data was pre-processed with median filtering and CLAHE. Then, the pre-processed images were classified with CNN and the parameters like hidden neuron count and epochs count were tuned using DOX and the main aim of the introduced method was to increase the accuracy and precision of the categorization phase. The accuracy of the proposed DOX-CNN was 15% enhanced than PSO-CNN, 12% enhanced than DHOA-CNN, 13% enhanced than SSA-CNN and 14% enhanced than EHO-CNN. Hence, the suggested brain tumour classification has attained high accuracy rate.

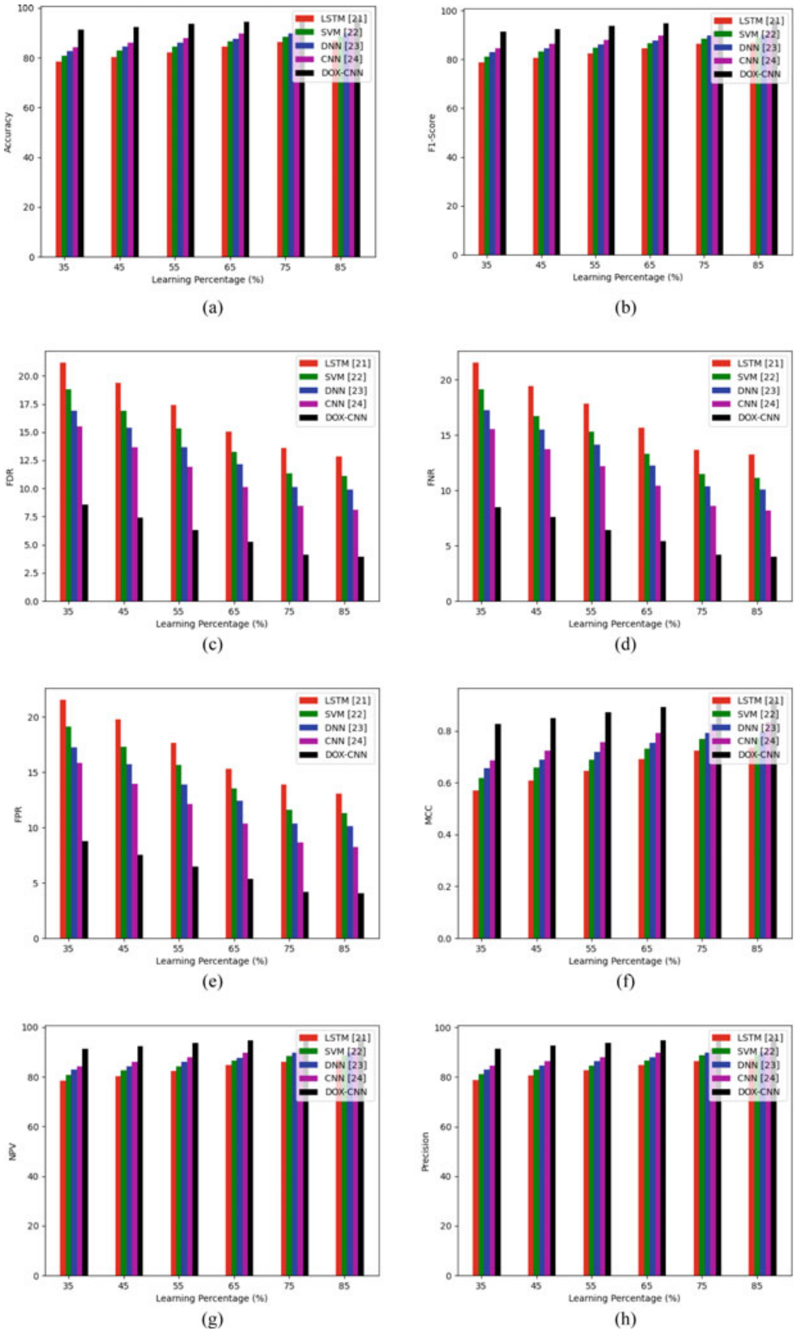


Fig. 5 Analysis of recommended brain tumour classification on traditional learning models at different learning rate concerning '**a** Accuracy, **b** F11-score, **c** FDR, **d** FNR, **e** FPR, **f** MCC, **g** NPV, **h** Precision, **i** Sensitivity, **j** Specificity'

Table 1 Analysis of introduced brain tumour classification method on algorithms

Measures	PSO-CNN [16]	DHOA-CNN [17]	SSA-CNN [18]	EHO-CNN [19]	DOX-CNN
"Accuracy"	0.902258	0.899516	0.919677	0.936613	0.958226
"Sensitivity"	0.903288	0.899457	0.919566	0.936163	0.958506
"Specificity"	0.901206	0.899576	0.919791	0.937072	0.957939
"Precision"	0.903288	0.901472	0.92133	0.93826	0.958812
"FPR"	0.098794	0.100424	0.080209	0.062928	0.042061
"FNR"	0.096712	0.100543	0.080434	0.063837	0.041494
"NPV"	0.901206	0.899576	0.919791	0.937072	0.957939
"FDR"	0.096712	0.098528	0.07867	0.06174	0.041188
"F1-Score"	0.903288	0.900463	0.920447	0.93721	0.958659

Table 2 Analysis of introduced brain tumour classification method on existing learning methods

Measures	LSTM [20]	SVM [21]	DNN [5]	CNN [22]	DOX-CNN
"Accuracy"	0.862258	0.884677	0.896452	0.913871	0.958226
"Sensitivity"	0.86339	0.885094	0.896266	0.91414	0.958506
"Specificity"	0.861102	0.884252	0.896642	0.913596	0.957939
"Precision"	0.863941	0.886509	0.89856	0.915308	0.958812
"FPR"	0.138898	0.115748	0.103358	0.086404	0.042061
"FNR"	0.13661	0.114906	0.103734	0.08586	0.041494
"NPV"	0.861102	0.884252	0.896642	0.913596	0.957939
"FDR"	0.136059	0.113491	0.10144	0.084692	0.041188
"F1-Score"	0.863665	0.885801	0.897411	0.914724	0.958659

References

1. Cha S (2006) Update on brain tumor imaging: from anatomy to physiology. Amer J Neuroradiol 27(3):475–487
2. Zulpe N, Pawar V (2012) GLCM textural features for brain tumor classification. Int J Comput Sci Issues 9:354–359
3. El-Dahshan EA et al (2014) Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. Expert Syst Appl 41(11):5526–5545
4. Liu M et al (2018) Anatomical landmark based deep feature representation for MR images in brain disease diagnosis. IEEE J Biomed Health Inf 22(5):1476–1485
5. Mohsen H, El-Dahshan ESA, El-Horbaty ESM (2018) Classification using deep learning neural networks for brain tumors. Future Comput Inform J 3(1):68–71
6. Cheng J et al (2015) Enhanced performance of brain tumor classification via tumor region augmentation and partition. PloS One 10(12)
7. Abbadi NKE, Kadhim NE (2017) Brain cancer classification based on features and artificial neural network. Int J Adv Res Comput Commun Eng 8(1):123–134

8. Ali S, Ismael A, Mohammed A, Hefny H (2020) Artificial intelligence in medicine an enhanced deep learning approach for brain cancer MRI images classification using residual networks. Artificial Intell Med 102
9. Anaraki AK, Ayati M, Kazemi F (2018) Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Integrative Med Res 39(1):63–74
10. Çinar A, Yildirim M (2020) Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. Med Hypotheses 139
11. Ge C, Gu IYH, Jakola AS, Yang J (2020) Enlarged training dataset by pairwise GANs for molecular-based brain tumor classification. IEEE Access 8:22560–22570
12. Habib H, Amin R, Ahmed B, Hannan A (2021) Hybrid algorithms for brain tumor segmentation, classification and feature extraction. J Ambient Intell Human Comput
13. Khairandish MO, Sharma M, Jain V, Chatterjee JM, Jhanjhi NZ (2021) A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. IRBM. Available online, June 2021
14. Kesav N, Jibukumar MG (2021) Efficient and low complex architecture for detection and classification of brain tumor using RCNN with two channel CNN. Available online, May 2021
15. Afshar P, Mohammadi A, Plataniotis KN (2020) BayesCap: a Bayesian approach to brain tumor classification using capsule networks. IEEE Signal Process Lett 27:2024–2028
16. Pedersen MEH, Chipperfield AJ (2010) Simplifying particle swarm optimization. Appl Soft Comput 10(2):618–628
17. Brammya G, Praveena S, Ninu Preetha NS, Ramya R, Rajakumar BR, Binu D (2019) Deer hunting optimization algorithm: a new nature-inspired meta- heuristic paradigm. 24 May 2019
18. Jain M, Singh V, Rani A (2019) A novel nature-inspired algorithm for optimization: squirrel search algorithm. Swarm Evol Comput 44:148–175
19. Yilmaz S, Sen S (2020) Electric fish optimization: a new heuristic algorithm inspired by electrolocation. Neural Computing and Applications 15
20. Amin J, Sharif M, Raza M, Saba T (2020) Brain tumor detection: a long short-term memory (LSTM)-based learning model. Neural Comput Appl 32:15965–15973
21. Rao CS, Karunakara K (2022) Efficient detection and classification of brain tumor using kernel-based SVM for MRI. Multimedia Tools and Applications 81:7393–7417
22. Devi S (2020) Performance prediction using deep learning technique in education sector. The International Journal of Analytical and Experimental Modal Analysis

Predicting the Presence of Lung Cancer Using Functional Networks

B. Usha Priya and V. Lokeswara Reddy

Abstract Predicting the lung cancer presence in human body is very important and also a daunting task to achieve. In the current study, a Functional Network (FN)–based model is developed to predict whether lung cancer exists in a human body. Hoarseness, anxiety, chronic illness, fatigue, allergies, wheezing, coughing, shortness of breath, swallowing difficulty and chest pain are identified to be the significant symptoms of lungs cancer. It has also been observed that the use of alcohol and/or tobacco also plays a significant role in the development of lungs cancer. Hence, all of these indicators are used as model inputs to in this study. The developed FN model is observed to have a high accuracy and success rate in predicting the presence of lungs cancer with a coefficient of determination (R^2) value of 0.93. This model would largely help the patients and medical staffs to identify the lungs cancer presence in an early stage, which may save many lives.

Keywords Lung Cancer · Cancer Diagnosis · Prediction Analysis · Artificial Intelligence · Functional Network

1 Introduction

The leading cause of death globally for both men and women is lung cancer [1, 2]. Other research indicates that in 2015, lung cancer made up around 13% of all cancer diagnoses in the United States. The American Cancer Society [3] estimates that 27% of all cancer-related fatalities are caused by lung cancer. Lung nodules that are still developing need to be carefully checked and tracked. Although various treatments are being used to treat lung cancer, the diagnosis of the same is quite poor since doctors can only detect the disease when it has advanced. In order to

B. Usha Priya (✉)
Department of CSE, JNTUA, Ananthapuramu, Andhra Pradesh, India
e-mail: ushapriya512@gmail.com

V. Lokeswara Reddy
Department of CSE, KSRM College of Engineering, Kadapa, Andhra Pradesh, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_55

easily reduce the mortality rate with efficient control, early prediction before the last stage is crucial. Even with the right treatment and diagnosis, the prognosis for lung cancer is quite encouraging. Lung cancer survival rates vary from person to person. Age, gender, and race are all factors, as well as health status. These days, artificial intelligence (AI) and machine learning (ML) is essential for the early diagnosis and prediction of medical disorders that keep people safe. Today, AI and ML have taken over the medical industry making the diagnosis procedure simple and deterministic. They make it simple to interpret genuine features or information, analyze data, and identify the root cause of illnesses. This in turn, assists doctors in determining the underlying causes of disorders, and aids the treating physicians in making a more accurate diagnosis of the illnesses, saving time and money.

The artificial intelligence (AI) techniques have also shown great efficiency in solving prediction problems in various fields of science and engineering. Hence, this study aims to investigate the ability of a novel AI technique, namely functional network (FN) to predict the presence of lungs cancer in a human body. FN is basically a recently introduced powerful alternative to the artificial neural networks (ANN). It has optimizing performance in giving accurate results in case of complex problems. For model training, required data sets are collected from residents, patients and medical staffs with the help of a questionnaire. Subsequently, the significant predictors of lungs cancer are identified and used to train the FN model. This model will largely help to predict the presence of lungs cancer in an early stage with simple information and symptoms of the patient. Thus, necessary treatments could be arranged at right time and many lives can be saved subsequently.

2 Literature Review

Researchers in the fields of oncology and artificial intelligence-based medical aid have both been quite concerned about lung cancer. While some researches have concentrated on early lung cancer diagnosis [4], some have built methods for the detection and diagnosis of lung cancer [5]. Fuzzy logic and neural networks have been used in studies regarding the diagnosis of lung cancer [6]. Hybrid neuro-fuzzy approaches have been employed in other investigations [7]. These techniques are unreliable because they cannot create a trustworthy medical diagnosis system with growing database sizes. Studies based on cutting-edge machine learning ideas, including decision trees [8], have shown improved dependability in comparison to those antiquated techniques. Prognostic models for Non-Small-Cell Lung Cancer (NSCLC) based on neural networks were introduced by Hanai and colleagues [9]. They used 125 NSCLC patients and 17 possible input risk variables to build their models. Research on the variability in smokers' risk of developing lung cancer based on a variety of characteristics was published by Kattan and Bach [10]. They evaluated the impact of various variables on the level of lung cancer risk. They discovered that although 0.8% of 51-year-old women who smoked a pack of cigarettes per day for

28 years had lung cancer, 15% of males over 68 who had smoked two packs per day for 50 years and kept smoking had the disease.

Based on data mining, Ramachandran et al. [11] developed an early detection method for lung cancer that made use of 11 key parameters. They carried out trials utilizing a database of 746 samples, but they made no reference to the source of their database. In order to determine the risk factor for lung cancer in 2014, Thangaraju and colleagues employed data mining techniques [6]. For grouping and classification, they employed Bayes Trees and Decision Tables. 303 samples were used in the trials. Based on 11 symptoms, Manikandan and colleagues created a hybrid neuro-fuzzy system for lung cancer prediction [8]. Using samples from a database of 271 people, they used 163 samples (221 medical situations and 50 normal persons). The symptoms that can be utilized to predict lung cancer were defined by Arulananth and Bharathi [12]. They distinguished between the characteristics that serve as diagnostic indicators and the cancer-related symptoms. They classified the diagnostic signs according to age, sex, family history of cancer, smoking, radiation, radon, chemical exposure, and air pollution exposure. On the other hand, they identified anorexia, a persistent cough, hemoptysis, chest discomfort, exhaustion, weight loss, chronic lung inflammation, and wheezing as signs that pointed to the existence of malignancy. Senthil and Ayshwaya [9] employed evolutionary algorithms and neural networks in 2018 to quantify the risk level of lung cancer based on risk variables. These algorithms were employed to analyze the UCI Global Lung Cancer Database, which had only 32 samples, and the utilized symptoms were general.

A clinical risk prediction model for lung cancer was recently developed by Markaki [13] and colleagues based on smoking symptoms. They were based on factors such as age, weight, height, hours spent in polluted areas, frequency of coughing, sex, number of years of smoking, daily cigarette use, and years after quitting smoking. Other research made use of cutting-edge machine learning techniques like random trees and random forests, which proved excellent at classifying large databases [14]. Others, however, have used radiation image processing methods to assess the presence or absence of lung cancer [15]. In other studies, the prediction of NSCLC patients' death in the American military healthcare system was the main topic [16]. Therefore, it was desirable to look for other characteristics in addition to smoking and age in order to construct a robust model for predicting lung cancer risk.

3 Functional Networks (FN)

The idea of FN is said to have revolutionized how complex systems have been studied during the past ten years [17]. Its architecture is based on the structure of the physical world [18–22]. A sample network structure consisting of several neurons and neuron functions is presented in Fig. 1a. In FN, these neuron functions are established entirely from the data analysis, making it more convenient to deal with complex problems. The initial architecture for the problem is also designed based on the data type and characteristics. Hence, FN is commonly called as a problem- or data-driven approach.

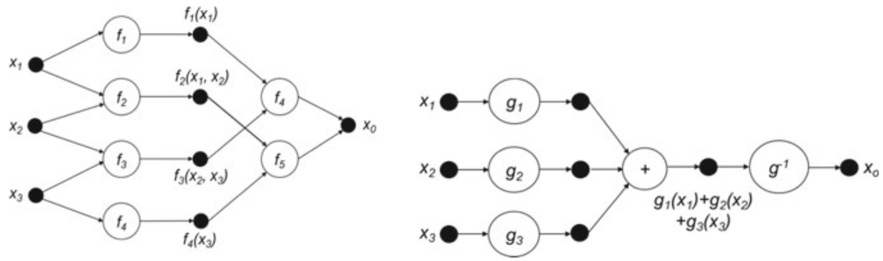


Fig. 1 A typical FN and its associativity FN

FN has also the ability to take into account the domain knowledge (i.e., properties of the function being modelled) while establishing the relationship among variables.

In this study, the most recent and simplest form of the FN is used to predict the presence of lungs cancer, called Associativity FN. This technique uses the theory of functional equations to convert any multi-input network with s inputs $(x_1, x_2,\ldots, x_s)$ and one output (x_{s+1}) to a simplified network [18, 19]. For instance, Fig. 1b represents the associativity FN of a general FN shown in Fig. 1a. The mathematical expression of an associative FN for a variable x is given in Eq. 1 below.

$$f(x) = \sum_{i=1}^{m} a_i \varphi_i(x) \tag{1}$$

where a_i = coefficient of $\phi(x)$ at ith degree, ϕ_i = shape function (SF) of variable x at degree i (which can be polynomial, exponential or logarithmic, etc.), and m = degree of SF.

By using above equation, the input functions of a problem can be summed up to predict the output function as:

$$f_{s+1}(x_{s+1}) = f_1(x_1) + f_2(x_2) + \cdots + f_s(x_s) \tag{2}$$

Accordingly, the error in the jth data (e_j) can be obtained as:

$$e_j = f_1\left(x_{1j}\right) + f_2\left(x_{2j}\right) + \cdots + f_s\left(x_{sj}\right) - f_{s+1}\left(x_{s+1,j}\right) \tag{3}$$

In order to estimate the values of a_i, the sum of squared error E_n for all n data sets is minimized as follows:

$$E_n = \sum_{j=1}^{n} e_j^2 = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_i \left[\varphi_i\left(x_{1j}\right) + \varphi_i\left(x_{2j}\right) + \ldots + \varphi_i\left(x_{sj}\right) - \varphi_i\left(x_{s+1,j}\right)\right] \right)^2 \tag{4}$$

subject to

$$f_k(x_0) = \sum_{i=1}^{m_k} a_{ki}\varphi_{ki}(x_0) = \alpha_k; k = 1, 2, ..., s+1 \tag{5}$$

where x_0 and α_k are constants.

Further, an auxiliary function E_λ is defined using Lagrangian multipliers as:

$$E_\lambda = \sum_{j=1}^{n}\left[\sum_{k=1}^{s+1}\sum_{i=1}^{m} a_{ki}\varphi_{ki}\left(x_{kj}\right)\right]^2 + \sum_{k=1}^{s+1}\lambda_k\left[\sum_{i=1}^{m} a_{ki}\varphi_{ki}(x_0) - \alpha_k\right] \tag{6}$$

The minimum value of E_λ can be obtained by using following equations:

$$\frac{\partial E_\lambda}{\partial a_{kr}} = 2 \times \sum_{j=1}^{n}\left[\sum_{k=1}^{s+1}\sum_{i=1}^{m} a_{ki}\varphi_{ki}\left(x_{kj}\right)\right] \times \varphi_{kr}\left(x_{jk}\right) + \lambda_k\varphi_{kr}(x_0) = 0; \quad k = 1, 2, ..., s+1 \text{ and } r = 1, 2, ..., m \tag{7}$$

$$\frac{\partial E_\lambda}{\partial \lambda_k} = \sum_{i=1}^{m} a_{ki}\varphi_{ki}(x_0) - \alpha_k = 0; k = 1, 2, ..., s+1 \tag{8}$$

The above systems of linear equations have $k \times (m+1)$ equations and $k \times (m+1)$ unknowns, which is solved to get the model coefficients a_{ki} ($k = 1, 2,\ldots, s+1$).

4 Data Collection

Data used in the study are collected by survey done in different cancer hospitals in the Hyderabad city of India. All the data collected are the real-time data of patients and nonpatients who had undergone all the necessary tests. In the survey, both patients and medical staffs were interacted and the details were filled very carefully and accurately. Approximately, 150 persons were interacted in this survey, and finally, the full details of 116 persons could be recorded successfully. Approximately 40% of females were included in this survey. Survey participants represented a good mix of age structure. However, the percentage of elderly was somewhat higher (approximately, 60%), as they are more affected by lungs cancer in India. Table 1 shows a sample of data sets collected in this study through face-to-face interactions.

Table 1 Sample data sets

Use of alcohol/ tobacco (x_1)	Anxiety (x_2)	Hoarseness (x_3)	Chronic disease (x_4)	Fatigue (x_5)	Wheezing (x_6)	Unexplained weight loss (x_7)	Coughing (x_8)	Shortness of breath (x_9)	Swallowing difficulty (x_{10})	Chest pain (x_{11})	Lung cancer
0	1	1	0	0	1	1	1	1	1	1	1
1	1	1	1	1	0	0	1	0	1	0	1
1	1	1	0	0	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	1	0	0	0
1	1	0	1	0	1	0	1	1	0	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	1	1	0	1	1	1	0	1	1	1	1
1	0	1	1	1	0	0	0	0	1	1	1
0	1	1	1	1	0	1	0	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0
1	1	1	1	1	1	0	1	1	1	0	1

5 Results and Discussion

The significant predictors of the presence of lungs cancer were identified with the help of Pearson's correlation analysis. The results (Pearson's correlation coefficient R and corresponding p-value) obtained from this analysis are presented in Table 2. It can be observed that the predictors are highly significant at the level of $p <$ 0.001. The multi-collinearity among these variables was also tested and the highest correlation coefficient was observed to be 0.358. This suggested that the variables are independent predictors of lungs cancer. Hence, all these parameters were used for training the FN model. Of the total data, 70% were used for model training and 30% were used for the model testing. The total data was divided into these two groups by covering the wide ranges of variables in each group. A MATLAB function "dividerand" was used to fulfil this requirement.

With the help of Eqs. (1) and (2), an associative FN model was built to predict the presence of lungs cancer (1 = yes, 0 = no).

$$\text{Cancer}_{\text{Presence}} = \sum_{j=1}^{s}\left(\sum_{i=1}^{m} a_{ji}\varphi_{ji}(x_j)\right) \tag{9}$$

where s is the number of input variables, m is the degree of SF, and a_0 is a constant term.

For developing the FN-based cancer prediction model, an appropriate SF was first selected along with its proper degree. The performance of different functions like polynomial, exponential, sin(.), cos(.) and tan(.) at different degree were tested. Finally, it was observed that all functions are able to produce a R^2-value of '1' between the predicted and perceived outputs at a degree of 5 and above. However, to keep the model structure simple but efficient, a lower degree of two was selected. At this degree, the polynomial function produced the best results with R^2-value of 0.93. Hence, the FN model was developed using polynomial function with degree two. The network structure of the model is presented in Fig. 2. Here, all input variables are fed in the input layer and the output is obtained in the last layer. The function f represents the polynomial function with degree two.

The FN model coefficients as obtained from the data analysis are presented in Table 3. Inputting these values in Eq. 9, the mathematical structure of the cancer prediction model is derived as follow:

$$\begin{aligned}\text{Cancer}_{\text{Presence}} = 0.0488 &+ 0.1475 \times x_1 + 0.0607 \times x_2 + 0.135 \times x_3 \\ &+ 0.0756 \times x_4^2 + 0.0293 \times x_5 + 0.0084 \times x_6 \\ &+ 0.370 \times x_7 + 0.0386 \times x_8 + 0.1152 \times x_9 \\ &+ 0.0847 \times x_{10}^2 + 0.0584 \times x_{11}\end{aligned} \tag{11}$$

where $x_1, x_2, \ldots, x_{11}$ are the terms as defined in Table 3.

The developed model returns a value in decimal, instead of exactly 0 or 1. Hence, a score < 0.5 could be understood as the absence of lungs cancer and a value > 0.5

Table 2 Significance of cancer predictors

Parameter	Use of alcohol/ tobacco (x_1)	Anxiety (x_2)	Hoarseness (x_3)	Chronic disease (x_4)	Fatigue (x_5)	Wheezing (x_6)	Unexplained weight loss (x_7)	Coughing (x_8)	Shortness of breath (x_9)	Swallowing difficulty (x_{10})	Chest pain (x_{11})
R	0.436	0.51	0.461	0.417	0.584	0.393	0.662	0.417	0.45	0.430	0.460
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

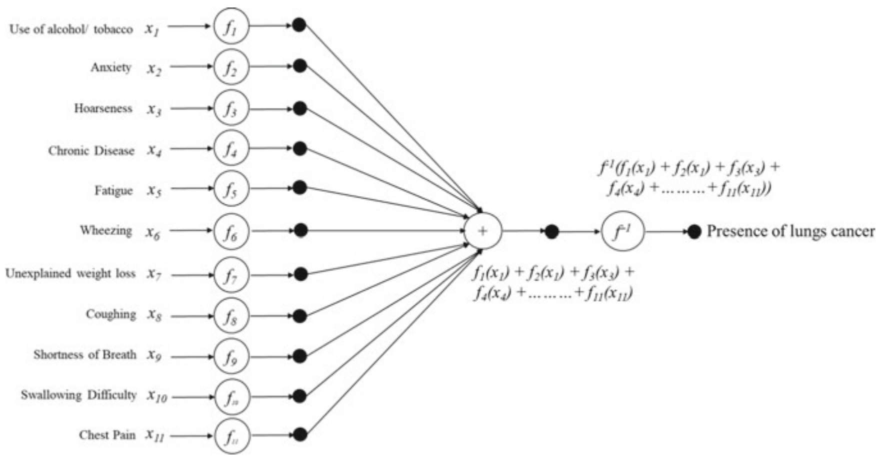


Fig. 2 Associativity FN of the lung cancer model

could be considered as the presence of it. The training and testing performances of the developed model is shown in Fig. 3, and prediction results in terms of different statistical parameters are shown in Table 4. As observed inn Fig. 3, the predicted values are very close to the ideal line of fit in both training and testing stages with R^2-values of 0.933 and 0.937, respectively. The model performance in terms of the root mean square error (RMSE), average absolute error (AAE) and maximum absolute error (MAE) in Table 4 also suggests that the developed model has a very high efficiency in predicting the presence of lungs cancer in a human body.

Table 3 FN model coefficients

Parameter	Constant	Use of alcohol/tobacco (x_1)	Anxiety (x_2)	Hoarseness (x_3)	Chronic disease (x_4)	Fatigue (x_5)	Wheezing (x_6)	Unexplained weight loss (x_7)	Coughing (x_8)	Shortness of breath (x_9)	Swallowing difficulty (x_{10})	Chest pain (x_{11})
Coefficient at degree 1	0.0488	0.1475	0.0607	0.135	0	0.0293	− 0.0084	0.3704	0.0386	0.1152	0	0.0584
Coefficient at degree 2		0	0	0	0.0756	0	0	0	0	0	0.0847	0

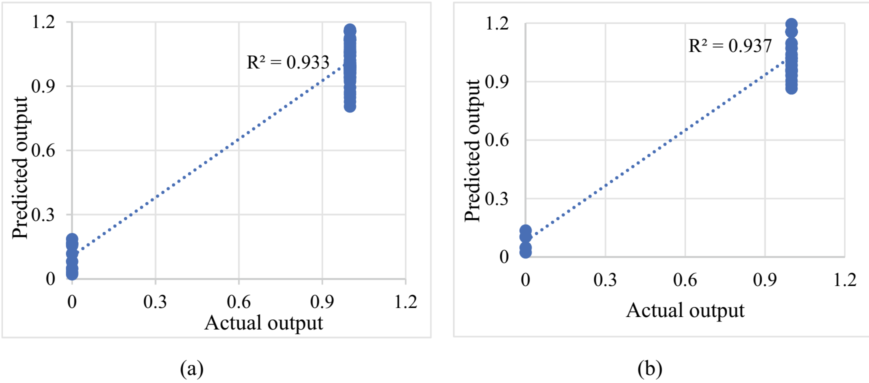


Fig. 3 Prediction performance of the developed model with **a** training and **b** testing data

Table 4 Prediction results of developed model

Parameter	Training stage	Testing stage
R^2	0.933	0.937
RMSE	0.094	0.095
AAE	0.074	0.078
MAE	0.195	0.195

6 Conclusion

This study has basically analyzed the real-world data to identify the significant predictors of lungs cancer in a human body. Subsequently, a novel AI technique namely, FN has been implemented to develop a mathematical model to predict the presence of lungs cancer in an early stage. Through a thorough investigation of human health, symptoms and unhealthy habits, this study has arrived at following conclusions:

1. The most notable symptoms of lungs cancer are hoarseness, anxiety, chronic illness, fatigue, allergies, wheezing, coughing, shortness of breath, difficulty swallowing, and chest pain.
2. The regular consumption of alcohol and/or tobacco also leads to lungs cancer.
3. Based on above factors, the prediction of the presence of lungs cancer can be successfully done in an early stage.
4. FN technique is highly efficient to predict the same with a high R^2-value of above 0.93 with averaged observations.
5. The model performance is also successful in terms of other statistical parameters like RMSE, AAE and MAE.

The outcomes of this study would largely help the public and medical staffs to detect the presence of lungs cancer promptly and provide necessary treatment on time. Future studies could also improvise the prediction results with the help of any other AI technique.

References

1. Nall R (2018) Medical news today. Available online at: https://www.medicalnewstoday.com/articles/323648.php. Accessed 14 March 2022
2. Brocken P, Kiers BA, Looijen-Salamon MG et al (20121) Timeliness of lung cancer diagnosis and treatment in a rapid outpatient diagnostic program with combined FDG-PET and contrast enhanced CT scanning. Lung Cancer 75(3):336–341
3. Vivekanandan P (2013) An efficient SVM based tumour classification with symmetry non-negative matrix factorization using gene expression data. 2013 international conference on information communication and embedded systems (ICICES). IEEE, USA, pp 761–768
4. Billah M, Islam N (2016) An early diagnosis system for predicting lung cancer risk using adaptive neuro fuzzy inference system and linear discriminant analysis. J MPE Mol Pathol Epidoemiol 1(3):1–4
5. Tiwari S, Walia N, Singh H, Sharma A (2015) Effective analysis of lung infection using fuzzy rules. Int J Bio-Sci Bio-Technol 7(6):85–96
6. Thangaraju P, Barkavi G, Karthikeyan T (2014) Mining lung cancer data for smokers and nonsmokers by using data mining techniques. Int J Adv Res Comput Commun Eng 3(7):7622–7626
7. Christopher T, Jamera J (2016) Study of classification algorithm for lung cancer prediction. Int J Innovat Sci Eng Technol 3(2):42–49
8. Manikandan T, Bharathi N, Sathish M, Asokan V (2017) Hybrid neuro-fuzzy system for prediction of lung diseases based on the observed symptom values. J Chem Pharmaceut Sci 8:69–76
9. Senthil S, Ayshwarya B (2018) Lung cancer prediction using feed forward back propagation neural networks with optimal features. Int J Appl Eng Res 13(1):318–325
10. Bach P, Kattan M, Thornquist M (2003) Variations in lung cancer risk among smokers. J Natl Cancer Inst 95(6):470–478
11. Ramachandran P, Girija N, Bhuvaneswari T (2014) Early detection and prevention of cancer using data mining techniques. Int J Comput Appl 97(13):48–53
12. Bharathi H, Arulananth TS (2017) A review of lung cancer prediction system using data mining techniques and self-organizing map (SOM). Int J Appl Eng Res 12(10):2190–2195
13. Markaki M, Tsamardinos I, Langhammer A, Lagani V, Hveem K, Røe OD (2018) A validated clinical risk prediction model for lung cancer in smokers of all ages and exposure types: a hunt study. EBioMedicine 31:36–46
14. Mishra A, Ratha B (2016) Study of random tree and random forest data mining algorithms for microarray data analysis. Int J Adv Elcetric Comput Eng 3(4):5–7
15. Saii M, Mia A (2015) Lung detection and segmentation using marker watershed and Laplacian filtering. Int J Biomed Eng Clin Sci 1(2):29–42
16. Jeong C, Jeong S, Hong S (2012) Nomograms to predict the pathological stage of clinically localized prostate cancer in Korean men: comparison with Western predictive tools using decision curve analysis. Int J Urol 19(9):846–852
17. Zanin M et al (2016) Combining complex networks and data mining: why and how. Phys Rep 635:1–44
18. Castillo E (1998) Functional networks. Neural Process Lett 7:151–159
19. Castillo E, Cobo A, Gomez-Nesterkin R et al (2000) A general framework for functional networks. Networks 35(1):70–82
20. Castillo E, Cobo A, Gutierrez JM et al (1998) An introduction to functional networks with applications. Kluwer Academic Publishers, Boston

21. Castillo E, Cobo A, Manuel J et al (2000) Functional networks: a new network-based methodology. Computer-Aided Civil and Infrastructure Engineering 15:90–106
22. Castillo E, Ruiz-Cobo R (1992) Functional equations in science and engineering. Marcel Dekker, New York

Lung Cancer Detection Through Deep Neural Networks Using CT Scan Images

S. Renu Deepti, B. Srivani, Ch. Kamala, and A. Sravani

Abstract Cancer is one of the most disastrous and life-threatening diseases to human beings among which lung cancer affects more than an estimated 2.3 million people around the world each year. Lung cancer is the main cause of mortality worldwide among all types of cancers. The suffering patient's survival rate can be improved by identifying the lung nodules accurately at a fast pace, that is, at an early stage. Nowadays, the field of automated diagnostic systems is becoming popular and thus most used in the diagnosis of any disease. To understand more about this, we can take an example of Image Processing which implements automated diagnostic system especially for the medical diagnosis is one such field where an automated diagnostic system. This will help in reducing the mortality rate and detecting the disease in the initial stage which can be considered as very remarkable in the bioinformatics field. In the majority of cases documented, patients were diagnosed when their disease had progressed to the point where there was no hope of a cure. To examine the presence of any symptoms or signs, screening is used. Hence, the main objective is to design and develop an Inception V3 algorithm to detect lung cancer that improves reliability.

Keywords Convolutional neural network · CT scan · Deep Learning · Flask framework

S. Renu Deepti (✉) · B. Srivani · Ch. Kamala · A. Sravani
IT Department, VNR VJIET, Hyderabad, India
e-mail: renudeepti_s@vnrvjiet.in

B. Srivani
e-mail: srivani_b@vnrvjiet.in

Ch. Kamala
e-mail: kamala_ch@vnrvjiet.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_56

1 Introduction

Cancerous cells can form tumors, immune system destruction, and other breakdowns that obstruct the body from handling them correctly. Lung cancer is a malignant tumor of the lungs that leads to irrepressible tissue cell proliferation. It is the world's most common disease, with 2,093,876 new cases reported in 2018. From 2005 to 2015, the research found that men's lung cancer rates fell 2.5% per year, while women's rates fell 1.2% per year. Symptoms incorporate a chronic cough, sputum striated with blood, chest pain, alterations in voice, increased conciseness of breath, and cyclical pneumonia or bronchitis, which normally do not appear until the malignancy has progressed [1]. One of the most valuable perilous components of lung cancer disease is still cigarette smoking which is most common in the world.

In the United States, burning weed is still responsible for 80% of accidents that are caused by lung cancer. Lung cancer is thought to be caused by exposure to radon gas, which is emitted through soil and construction materials [2]. For much better picture recognition which leads to accurate predictions, we are employing the InceptionV3 convolutional neural network architecture. Our major goal is to use attributes and information to diagnose the presence of lung cancer cells. The main things that should be considered while diagnosing to check for lung cancer are the attributes. The research checks for the feasibility of using an Artificial Neural Network model to identify lung cancer in an individual's body. The purpose of this research is to identify some significant factors that can cause lung cancer, make a neural network model that can be used for detection of lung cancer and to also know the cancer stage if it is present like malign, benign, or normal.

The main aim to be achieved from cancer screening is to decrease the count of demises caused by lung cancer or to get rid of cancer deaths completely. Screening is being applied to have benefits in clinical practice instead of in a research trial setting. False positives, cost, unintentional results, radiation exposure, and overdiagnosis are all issues that need to be addressed. Despite the truth that CT scan imaging is a substantial image depicting tool in the medical domain, it is difficult for clinicians to interpret and identify cancer from CT scan images, and visual elucidation of these CT scan images may be a fallible task that causes lung cancer detection to be delayed. As a result, clinicians may find computer-assisted diagnostics useful in properly identifying malignant cells.

2 Related Works

Till now, many have tried to develop a neural network frequently aiming at increasing the accuracy of lung cancer diagnosis.

Bhatia et al. in [3] have performed preprocessing on those images to get the areas of the lungs that are sensitive to cancer. Feature extraction was done using UNet and resNet models. Then used a combination of XGBoost and random forest which was provided with extracted features for predicting cancer using the LIDC-IRDI database, the accuracy obtained was 84%. Makaju et al. in [4] used Watershed segmentation for detection of nodules and SVM for the purpose of classification of nodules into Malignant or benign. He improved the image quality by calculating the weighted mean function. He used the Lung Image database consortium as a database which gave an accuracy of 92%. Faisal et al. in [5] worked with a number of classifiers such as MLP, Neural Network, Decision Trees, Naive Bayes, and SVM with a database obtained from UCI. From all the examinations, he found Gradient-boosted tree works well repository which gave an accuracy of 90%. Alakwaa et al. in [6] used a CT scan dataset from DSB and implemented the CNN algorithm. He predicted sensitive areas using UNet architecture. His model gave an accuracy of 86.6%.

Abdillah et al. in [7, 8] used CT scan Images which are from VIA and ELCAP databases. Region growing, Marker Controlled Watershed, and Marker Controlled Watershed with masking are applied. And among these, one with highest accuracy and robustness was Watershed with masking method. Alam et al. in [9] the preprocessing was done by image enhancement, segmentation, Image scaling, color space transformation, and contrast enhancement. He used a multi-classifier with the lung cancer dataset used for training which was taken from the UCI machine learning database. The precision was 97% for cancer identification and 87% for cancer prediction.

3 Problem Statement

The development of a Neural Network-based Detection Model, capable of identifying tumors of Lung Cancer in CT Scan Images of an individual and thereby categorizing them as benign, malignant, or none, while being completely automated in its functioning and requiring minimal human intervention from the start to finish.

4 Existing System

Currently, there exist the following types of testing methodologies, which are deployed at scale to detect LUNG CANCER:

1. **Computed tomography (CT) scan**—It is not like x-rays; it takes several pictures of the body and the computerized system combines those pictures as a part of the body that will be studied/analyzed. It is used to identify nodes/lumps present in the body
2. **Magnetic resonance imaging (MRI) scan**—MRI scans, like CT scans has radio waves in the place of x-rays.,. These scans are used to analyze lung cancer metastasis to the brain or spinal cord. These were used on a large scale at a point in time [10].

3. **Sputum cytology**—In the lab the mucus you cough up from your lungs is tested to find if it contains cancerous cells. The regular approach to do this is by collecting the samples three times each day for three consecutive days. Squamous cell lung tumors are more likely to be found by this because they start in the lung's primary airways.
4. **Needle biopsy**—A needle is regularly used to collect a sample from a mass. The disadvantage is that they only collect a little amount of tissue, which in a few situations may not be sufficient to do the diagnosis and for additional testing on cancer cells to assist doctors in choosing anticancer medications.

The challenges faced by the traditional methods are:

1. **Long turnaround times**—These results of the model can be obtained on the same day or in 1–2 days in a few situations, although the standard method gives us the result taking time such as 1–2 weeks or so.
2. **False-negative results—**Because there are no exact criteria for negative screening LDCT findings, sensitivity is usually ascertained by classifying new instances of lung cancer and giving the result as false negatives that appear under a year of a screening study. In the six trials that evaluated this characteristic, the sensitivity of LDCT for the identification of lung cancer extended from 80 to 100%.
3. **Radiation Exposures**—For the conventional/standard methods, the person must undergo many radiations. These methods have shown a huge increase in radiation exposure. The diagnosis of the chest/lungs for CT accounted for the major ratio of radiation exposure in cases of screening of lungs. Patients have to undergo/ recieve on an average of eight mSv over the treatment that has been undergone for almost 2–3 years, this involves both the detection and diagnostic/treatment assessment.

5 Proposed Method

Speedy recognition of lung cancer has flatter censorious, and image processing and deep learning techniques have made it possible. A lung cancer diagnosis has become a lot easier with deep learning. Lung patient Computer Tomography (CT) scan pictures were beneficial in this investigation to locate and classify lung nodules, as well as to demonstrate their malignancy stage. In this research, we are utilizing a deep learning CNN algorithm to detect lung cancer from CT-SCAN pictures, and to train CNN, we're using the CT-SCAN image dataset. The greater motivation of this research is to see how efficient classification algorithms are at detecting lung cancer early.

Deep neural networks have the benefit of a variable weight-sharing mechanism that improves the algorithm's performance. As a result, we set out to create a reliable diagnostic LUNG CANCER detection algorithm based on CT scan pictures. Our model seeks to use CNN for feature extraction, followed by Deep Learning InceptionV3 Algorithms to accurately characterize a CT scan image as belonging

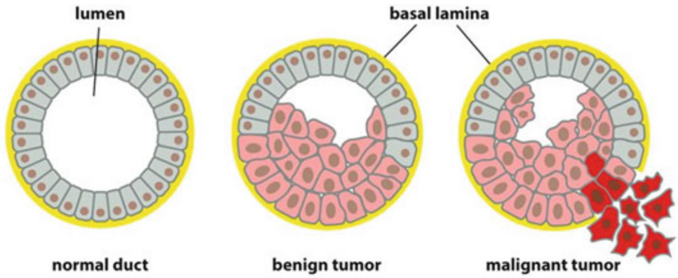


Fig. 1 Normal versus Benign versus Malignant Tumor

to an infected or healthy person. Deep learning is capable of learning a partial deep model on a partition of the overall data [11]. Data flow graphs allow us to understand more easily about a model. These are provided by a popular library called Tensorflow which is quite an intelligent toolset. It is capable of allowing programmers to construct wide-ranging neural networks with several layers.

To preprocess the photos, scale them to a given width, and do data augmentation, use Tensorflow's Keras preprocess function. The different information collectors utilize their own schemata for recording the data and the characteristics of different applications result in various data representations [12]. The Softmax function is utilized in the output layer of neural network models that foresee a multinomial probability distribution as an activation function. Softmax is acquired as the activation function for multi-class categorization that requires class membership on more than two outputs depicted in Fig. 1.

1. Ability to examine the deportment and efficiency of diverse visual models, from inception to Neural Architecture Search (NAS) networks, and then fine-tune them appropriately.
2. We will determine the model's classification ability by measuring the area lower than the graph (AUC) of a recipient operator curve. Visually analyze the behavior of these models by representing class activation maps (CAMs) or heatmaps for all the individual networks (ROC).
3. Ability to correctly diagnose lung cancer cases with an accuracy $\geq 90\%$.
4. Implementation of a fully automated lung cancer diagnosis system.

The following are the step-by-step procedure to build the model:

1. **Data Collection**: We have downloaded the dataset from a website named Kaggle.of CT Scan images.
2. **Data Preprocessing**: Resizing, flipping, zooming, and rotating are the few preprocessing techniques used.
3. **Building the model**: InceptionV3 is the model we're developing to detect lung cancer. In this case, every layer's output is utilized as an input to the next layers. We have used three techniques of regularization, namely L1, L2, and dropouts for building the best fit model.

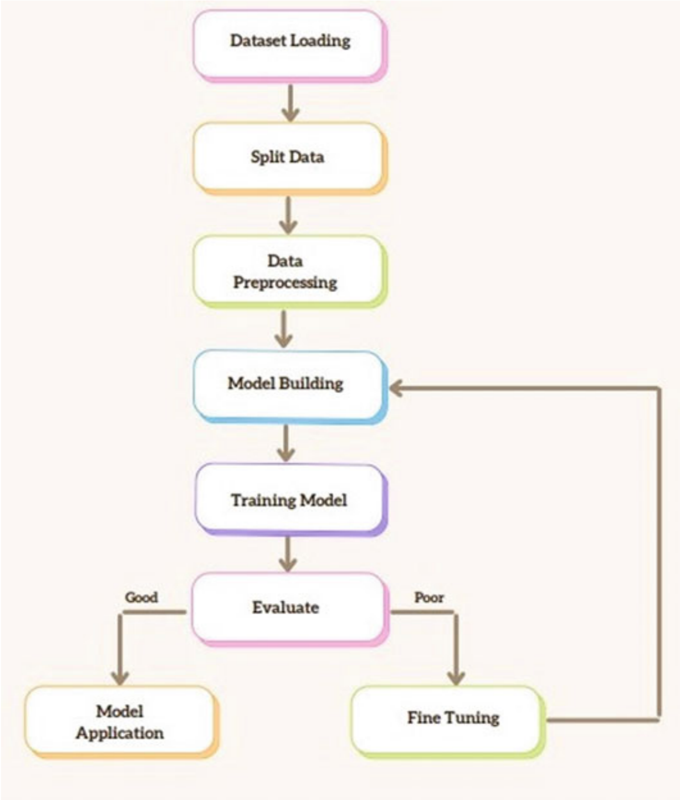


Fig. 2 Model workflow

4. **Evaluating the model**: We fine-tuned the model by changing multiple hyperparameters such as the count of neurons, activation function, optimizer, learning rate, size of the batch, and epochs to get better accuracy and loss curves. The model workflow is represented in Fig. 2.

6 Experimental Results

6.1 Dataset Description

A total of 1093 CT scans were included in the data set, of which 120 belonged to benign cases, 557 belonged to malignant cases, and 416 to normal cases. The current Model is trained in a way that uses 764 images as training data, 164 images as validation data, and 165 images as testing data. Uniformity between all cases is maintained in all three datasets. All these images were gathered from Open source

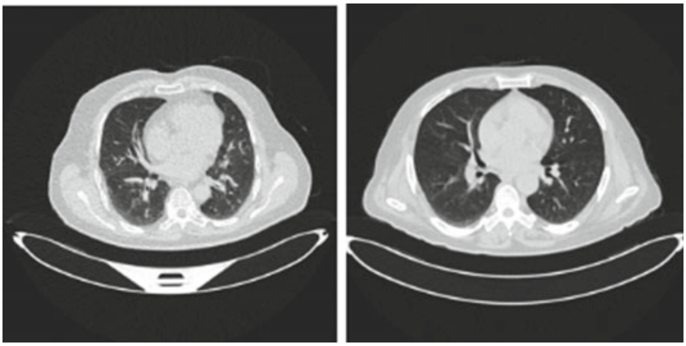


Fig. 3 CT scans of benign and malign cases

and are given due credit. The CT scans of selected benign and malignant cases and their counterparts with probable illness are shown in Fig. 3.

There is a few parameters and hyper parameters that need to be considered to build the best fit model. In the process, we have experimented with different optimizers, learning rates (0.1, 0.01, 0.2), and activation functions (relu, sigmoid, softmax, swish). The model has been trained on different optimizers to capture the suitable optimizer for our model. The table represents the accuracies obtained with different optimizers shown in Fig. 4.

So, the final model was built using an SGD optimizer with 0.001 learning rate, relu, and softmax activation functions. The model was run for 50 epochs over training and validation datasets resulting in the following accuracy and loss curves are represented in Fig. 5.

Figure 6 shows the classification result. Hence, the overall accuracy is 95.75, 97% for training, and 95.12% for the testing set. We have built a webpage using Flask which is a lightweight Python web framework. The main aim of building it is that everyone can use it without any difficulty and predict cancer easily and quickly.

This works just by uploading an individual's CT scan image of anyone to predict whether that person has a tumor or not. The three expected results are benign (non-cancerous tumor), malignant (cancerous tumor), and normal (no tumor). The output is obtained in seconds which helps the doctors or the individual who is using. This is very user-friendly, and everyone can easily understand it.

Optimizer	Training Accuracy	Validation Accuracy
Adam	86%	74.63%
RMSprop	77%	71%
SGD (0.001 LR)	97%	95.21%

Fig. 4 Optimizer function variations

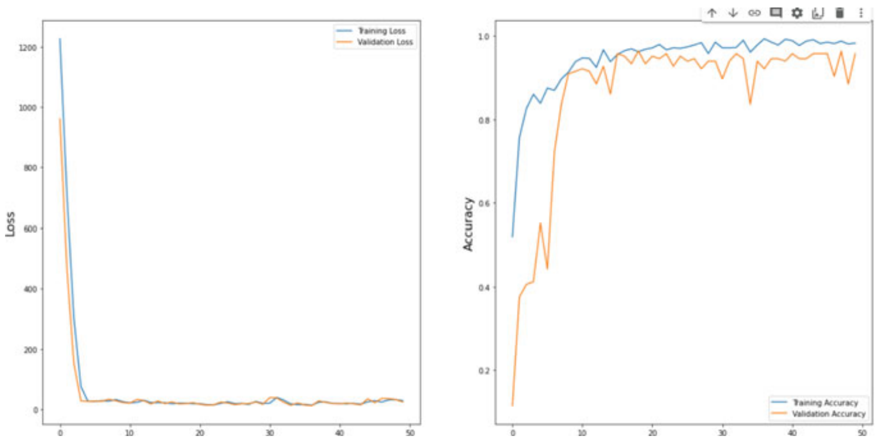


Fig. 5 Loss and accuracy curves

	Precision	Recall	F1-Score	Support
Bengin	0.76	0.89	0.82	18
Malignant	1.00	1.00	1.00	84
Normal	0.97	0.92	0.94	63
Accuracy			0.96	165
Macro avg	0.91	0.94	0.92	165
Weighted avg	0.96	0.96	0.96	165
Total Accuracy is 0.9575757575757575				

Fig. 6 Classification report

7 Conclusion and Future Work

We used InceptionV3 architecture for accurate image recognition and obtained an accuracy of 95%. Also, a model is implemented which can be accessed by anyone to easily predict lung cancer. In conclusion, our work is at a stage of minute improvements but a stabilized and working state with maximum accuracy. There are a few improvements like increasing the accuracy of the model by using ensembling algorithms, creating an application for lung cancer detection, and showing the size of the tumor and degree (stage) of cancer.

References

1. Deepti SR, Srivani B (2016) Efficient algorithm for mining high utility itemsets from large datasets using vertical approach. IOSR J Comput Eng 18(4):68–74

2. Siegel KD, Miller JA (2017) Cancer statistics. Cancer Journal for Clinicians 67(1):7–30
3. Bhatia S, Sinha Y, Goe L (2019) Lung cancer detection: a deep learning approach. Advances in Intelligent Systems and Computing 817
4. Makajua S, Prasad PWC, Alsadoona A, Singh AK, Elchouemi A (2018) 6th international conference on smart computing and communications, ICSCC 2017, 125, 107–114
5. Faisal MI, Bashir S, Khan ZS, Khan FH (2018) An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. In: 3rd international conference on emerging trends in engineering, sciences and technology
6. Alakwaa W, Nassef M, Badr A (2017) Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). International Journal of Advanced Computer Science and Applications 8(8)
7. Abdillah B, Bustamam A, Sarwinda D (2017) Image processing based detection of lung cancer on CT scan images. J Phys: Conf Ser 893(1):012063
8. Sharma D, Jindal G (2011) Computer aided diagnosis system for detection of lung cancer in CT scan images. International Journal of Computer and Electrical Engineering 3(5):714–718
9. Alam J, Alam S, Hossan A (2018) Multistage lung cancer detection and prediction using multi-class SVM classifier. In: International conference on computer, communication, chemical, material and electronic engineering
10. Bhatnagar D, Tiwari AK, Vijayarajan V, Krishnamoorthy A (2017) Classification of normal and abnormal images of lung cancer. IOP Conference Series: Materials Science and Engineering 263
11. Srivani B, Sandhya N, Padmaja Rani B (2020) Literature review and analysis on big DataStream classification techniques. Int J Knowledge Based Intelligent Eng Syst 24(3):205–215
12. Srivani B, Sandhya N, Padmaja Rani B (2021) An effective model for handling the big data streams based on the optimization enabled spark framework. Intelligent system design. Springer, Singapore, pp 673–696
13. Fan DP, Zhou T, Ji GP, Yi Z, Chen G et al (2020) Inf-net: automatic lung infection segmentation from CT images. IEEE Trans Med Imaging 39(8):2626–2637
14. Rahimzadeh M, Attar A (2020) A modified deep convolutional neural network for detecting pneumonia from chest X-ray images based on the concatenation of exception and resnet50v2. Informatics in Medicine Unlocked 19:100360
15. Li L, Qin L, Xu Z, Yin Y, Wang X et al (2020) Artificial intelligence distinguishes lung cancer from community-acquired pneumonia on chest CT. Radiology 296(2):E65–E71
16. Lortet-Tieulent J, Soerjomataram I, Ferlay J, Rutherford M, Weiderpass E, Bray F (2014) International trends in lung cancer incidence by histological subtype: adenocarcinoma stabilizing in men but still increasing in women. Lung Cancer 84(1):13–22

A Best Price Web Scraping Application for E-commerce Websites

Vinay Kumar Sriperambuduri, V. Sireesha, and Nagaratna P. Hegde

Abstract Web scraping is an approach for extracting loads of data from websites and handling this data as per need, for instance, saving to a local file on a computer or to a database in table or spreadsheet format. With the evolution of technology, there has been an increased demand of e-commerce websites, making online shopping the choice of every individual as it is easy and saves time. But at the same time plethora of choices have also caused confusion among buyers as to which product to choose based on the different prices offered for each product. Searching on several shopping portals and comparing is a tedious job. So, to make tasks easy, a web scrapping application for e-commerce websites has been developed and is presented in this paper. This application helps to collect data from several shopping portals, like Amazon, Flipkart, Myntra so on and displays them in a single catalog, so that the user can save time by avoiding browsing individual website to select the product that cost the user less. Based on the search query made by the customer, this application scrapes through various websites and show cases the prices offered by various vendors and shopping portals at one place.

Keywords Web Scraping · Cheerio · Puppeteer · HTTP request/response

V. K. Sriperambuduri (✉) · V. Sireesha · N. P. Hegde
Vasavi College of Engineering, Hyderabad, India
e-mail: s.vinaykumar@staff.vce.ac.in

V. Sireesha
e-mail: v.sireesha@staff.vce.ac.in

N. P. Hegde
e-mail: nagaratnaph@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_57

1 Introduction

Web scraping [1–3] is a process of extracting valuable and interesting text information from web pages. Web scraping can also be termed as web harvesting or web data extraction. This process could be bifurcated into three steps. The first step is to extract tons of unstructured data, which is held between the HTML elements of a web page. The second step is to annotate the data to provide it a semantic structure. The third step is to store a structured and semantic oriented data into files, excel sheets or database. This chapter focuses on step one and two, whereas step three is for future enhancement. One can manually extract the data, that is, by viewing it and by making an entry of that data for reference purpose. But this is not only a time-consuming task but also mundane and to reduce manual efforts we prefer to use automated processes implemented using a bot or web crawler. It is a form of copying, in which specific data is gathered and copied from the web, typically into a central local database or spreadsheet, for later retrieval or analysis.

A web scraping application will automatically gather data by copying it from multiple pages of websites based on a specific requirement for which it is being accumulated. It is built scalable enough to work for any kind of website and store the data being collected in easily readable format if one scrapes data for ethical purposes and does not aim to get access to confidential public data.

Websites being browsed by a user are only capable to display data to the user and do not bare the functionality to capture the data for future processing purpose. The user is then left with the option to manually copy and paste the data which is a very tedious job and can take many hours or sometimes days to complete. This manual effort can be replaced by making use of a web extracting software or an application.

The purpose for developing this best price application is for carrying out price data extraction by setting up custom web crawlers to fetch the product data from competitors' E-commerce portals for various brands. For a single crawl done, the product data extracted would contain various fields like web site name, brand name, product name, it's properties and the price. The product data would be unstructured and might not be fit for being used right away and so it can be run through a normalization system to do away with the unnecessary characters and parsing it for semantic analysis and by this we can ensure to avoid any problems that might occur due to special characters, symbols and numbers present in the raw data.

This chapter mainly focuses on the system developed to scrape data from various websites to display it at one place to save time and efforts of users without them having to go to different e-commerce websites to get product's prices and compare them. Literature survey included in Sect. 2, Proposed system included in Sect. 3, Implementation details included in Sect. 4 and Results and Discussion in Sect. 5.

2 Related Work

Web scraping [4–7] contributes to the expansion and upliftment of any business dealing with web data. New tools such as Puppeteer make it possible to scrape virtually anything. Several frameworks [8–16], libraries and APIs make it easier for a developer to perform web data extraction for ethical purpose. Some of the open-source tools used for web scraping are mentioned below:

1. Jsoup, an open-source Java-based tool uses the Document Object Mapping (DOM) method to extract HTML data and manipulate as per the business use case,
2. Puppeteer tool, for secured usage and is used as a NodeJS module,
3. Selenium API uses JavaScript and can easily be accessed from many other programming languages,
4. Scrapy, a framework in Python and is an open source,
5. BeautifulSoup, a python library for making HTTP calls and could be used for HTML extracting, etc.

   - Stock Market Statistics [17]: A company that want to analyze the price of the product in the market, compare it's past and current sales trend, and so would scrape tons of data from web to build the statistics reports.
   - Escalation of Business: To create new products and innovate faster, companies choose to scrawl through a vast amount of data accessible to the world. Take, for example, a job portal, for them to scale up the business, not only would they scrape data from websites like LinkedIn to look for professionals looking for jobs but also from company websites for job openings and makes it easier for companies to hire and for job seekers to view several job openings. This is a lucrative approach as the owners of job portal would get perks from the recruiters for advertising their job openings.
   - Data for content creation [14]: Companies can use web scraping for generation of leads and guides by accumulating information about potential customers and clients.
   - Social Media Analysis, Machine learning and large datasets: Data from social media websites can be gathered and analyzed. For instance, a television channel can scrape data posted on social media web sites like Twitter, Facebook, etc. for the shows being telecasted on their channel [16]. The large amount of raw data collected can be processed to generate reports on the show's rating and popularity among the public.
   - Brand surveying: Extracting data across various websites to study data about product brands for comparing ratings and prices offered by different vendors and online stores [12].

3 Proposed System

Collecting information about anything and everything present on a public domain is a common trend in the present time and this information is unstructured, further requiring to be deciphered or decoded to study its pattern and get some valuable information or report out of it. Books, article, encyclopedia, newspapers, magazines, catalogs etc. are some of the sources which are rich in content, but sources available on web are highly preferred because they contain updated information compared to any other source. An application developer can code a simple scraping application that can iteratively go through web pages within a predefined depth and get back all the necessary content being displayed on the page and then use it for doing various sort of analytics and research [17].

Websites can be penetrated by using the right tools as they follow rigid defense mechanism and firewalls to avoid easy access to hackers. As mentioned in Section II, there are various libraries, packages available and permitted to access the pages directly for ethical purposes.

Algorithm

Step 1: Fetching a webpage. Data like product name, vendor, and cost present in websites like Amazon, Flipkart will be extracted.

Step 2: Extracting data from a webpage. Using the request modules, we can fetch the webpage, and cheerio library provides us with utility functions to extract content on the page easily.

Step 3: Showcasing extracted content in a user-friendly manner. Data being scraped is stored in arrays and by looping through these arrays we retrieve this data and display it to the user.

Step 4: Capturing data in database, csv files, etc. If there be a need to create analytical reports for studying pricing trend, then this data can be fetched from the database.

Architecture

The proposed system adopted the following architecture shown in Fig. 1.

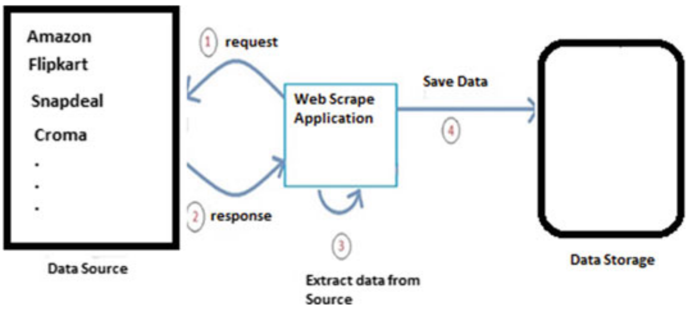


Fig. 1 Architecture

4 Implementation

The system cited in this paper uses HTML, CSS for front end and uses libraries of Node JS for back-end coding to connect to various websites and read the data by pointing to the tag name that can be captured using inspect properties.

1. Requirements:

   a. request-promise: Establishing a connection to the website by using HTTP calls. HTTP calls can be made effortlessly by using the Request module, which is a simple HTTP client.
   b. cheerio — jQuery for Node.js: Cheerio makes it easy to select, edit, and view DOM elements. Cheerio is a package provided by Node.js which follows a similar implementation as jQuery. It provides an API for parsing DOM data and manipulating that data.

2. Setting up the system repository:

   a. Creating a new project folder. Within that folder create an index.js file and home.js file. We'll need to install and require our dependencies. Opening command line to install request, request-promise, and cheerio packages.

3. The request is set up:

   a. Deciding on the website of which data needs to be extracted. The URL of this website needs to be passed to request to establish the connection via HTTP calls

4. The request being made:

   a. The request can now be made, once the above set up is successful.

5. Usage of the Data with the help of cheerio package:

   a. Use Cheerio's selector implementation makes it easy for a developer to code the logic with ease and extract necessary data.

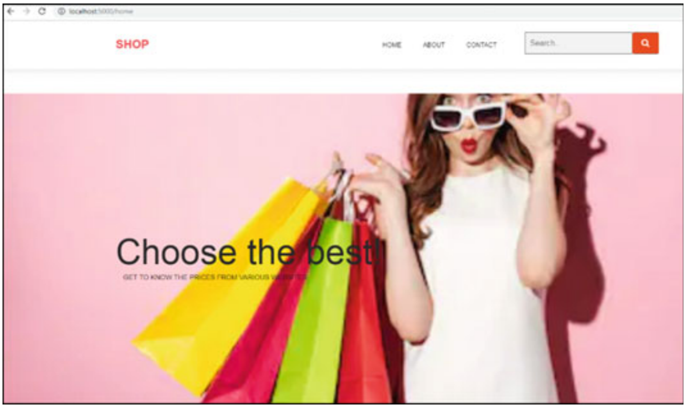


Fig. 2 home.ejs

6. Developer Tools used for getting the HTML tag names of the data to be scraped:
   a. Using inspector panel to find the element/class/ID names of the HTML constructs present on the web page.

5 Experimental Results

The web scraping application developed will be able to iteratively go through web pages within a predefined depth, this is done when the user enters the search keyword in the search field provided as shown in Fig. 2. The scraped content from different sources is then displayed on webpages as shown in Figs. 3, 4, 5 for the search keyword 'shoes'. The results show the products extracted from vendors along with the price.

6 Conclusion and Future Work

The proposed system explained in this paper is helpful to ensure service to users to make their daily life better, easier, and happier. This proposed system was developed using Node JS libraries and cheerio, puppeteer features which are available for creating web scrapping applications. The process of developing this application helped to explore several other aspects in the field of our study. This work provides a baseline approach in several ways using various tools for web scraping. Further, the advantage is that ethical web scraping offers to build large datasets, being accumulated daily, with ease by saving time, effort, and money.

In the future, we will improve this web application by extending our work in Travel and Tourism, Banking and Insurance and build large datasets in Stock trading.

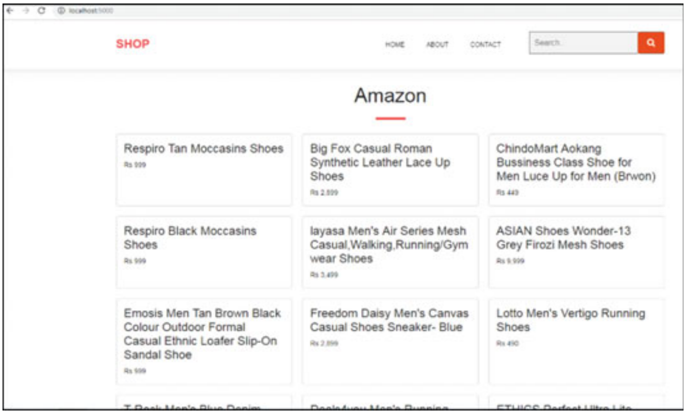


Fig. 3 Result data page1

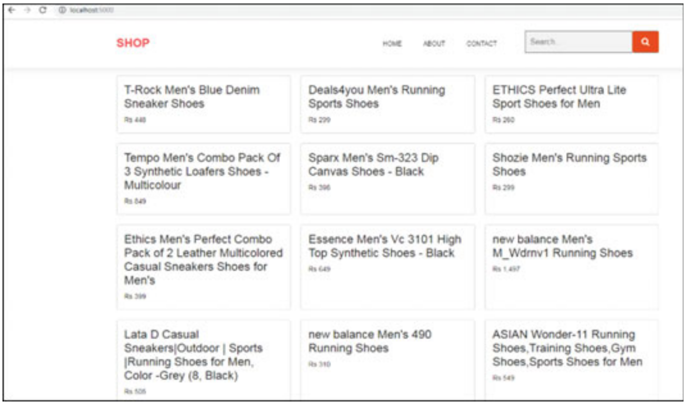


Fig. 4 Result data page2

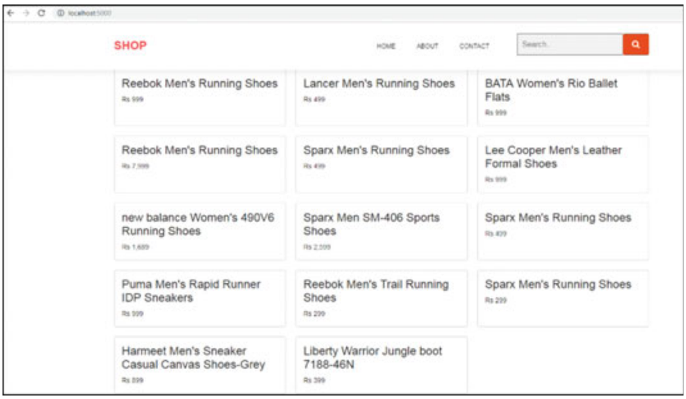


Fig. 5 Result data page3

Acknowledgements We thank Vasavi College of Engineering (Autonomous), Hyderabad for the support extended towards this work.

References

1. Uzun E (2020) A novel web scraping approach using the additional information obtained from web pages. IEEE Access 8:61726–61740. https://doi.org/10.1109/ACCESS.2020.2984503
2. Gao P, Han H, Guo J, Saeki M (2018) Stable web scraping: an approach based on neighbour zone and path similarity of page elements. Int J Web Eng Technol 13(4):301–333
3. Saurkar AV, Pathare KG, Gode SA (2018) An overview on web scraping techniques and tools. International Journal on Future Revolution in Computer Science & Communication Engineering 4(4):363–367. ISSN: 2454-4248
4. Sabria IAA, Mana M, Bakarb WAWA, Roseb ANM (2019) Web data extraction approach for deep web using WEIDJ. In: 16th international learning & technology conference 2019. ScienceDirect, Procedia Computer Science, vol 163, pp 417–426
5. Singrodia V, Mitra A, Paul S (2019) A review on web scrapping and its applications. IEEE INSPEC Accession Number: 18959922
6. Schintler LA, McNeely CL (2018) Web scraping. Encyclopedia of big data. Springer International, pp 1–3
7. Diouf R, NgorSarr E, Sall O, Birregah B, Bousso M, Mbaye SN (2019) Web scraping: state-of-the-art and areas of application. In: IEEE international conference on big data (Big Data)
8. Thomas DM, Mathur S (2019) Data analysis by web scraping using python. In: 3rd international conference on electronics, communication and aerospace technology (ICECA)
9. Bradley A, James RJE (2019) Web scraping using R. Adv Methods Pract Psychol Sci. https://doi.org/10.1177/2515245919859535
10. Himawan A, Priadana A, Murdiyanto A (2020) Implementation of web scraping to build a web-based instagram account data downloader application. International Journal on Informatics for Development (IJID)
11. Parvez MS, Tasneem KSA, Rajendra SS, Bodke KR (2018) Analysis of different web data extraction techniques. In: International conference on smart city and emerging technology (ICSCET)
12. Asif AA, Khandkar M, Rasel Turag SM (2019) A comparative study and integration of multiple ecommerce websites. Daffodil International University
13. Karthikeyan T, Karthik S, Ranjith D, Vinoth Kumar V, Balajee JM (2020) Personalized content extraction and text classification using effective web scraping techniques. International Journal of Web Portals (IJWP) 11(2)
14. Salah M, Okush BA, Rifaee MA (2019) A comparison of web data extraction techniques. IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)
15. Gunawan R, Rahmatulloh A, Darmawan I, Firdaus F (2018) Comparison of web scraping techniques: regular expression, HTML DOM and Xpath. In: Proceedings of the 2018 international conference on industrial enterprise and system engineering (IcoIESE 2018)
16. Parikh K, Singh D, Yadav D, Rathod M (2018) Detection of web scraping using machine learning. IEEE ISSN (Online) 2456-3293
17. Oancea B, Necula M (2019) Web scraping techniques for price statistics—the Romanian experience. Computer Science, Statistical Journal of the IAOS2019

GAN-Based Handwritten Digit Recognition System

P. Annan Naidu, Sasibhusana Rao Pappu, and Sura Papa Rao

Abstract Machine learning and computer vision play a vital role in real-time applications such as in industries, marketing, education, finance, automative, security, manufacturing and health. Machine learning process includes understanding the data without being explicitly programmed, and computer vision process includes understanding image features from various sensing devices for further image processing with its extensive algorithms or deep learning techniques Computer applications gone intense transformation from simple data processing to complex problem solving aspects, and this may be possible due to the huge availability of data from different data sources. In this connection, machine learning adds the additional advantages to the computer applications with its diversity of models. In the current scenario, the research has been more focused on the latest computing technologies like computer vision, deep learning, artificial intelligence and data science. The main objective of this study is to elegant the various aspects of machine learning and computer vision with its applications to transform the changes in the field of computer applications. As an application concern, this paper also discussed the importance of Generative Adversarial Network (GAN) and GAN Lab implementation with handwritten digit images.

Keywords Discriminator · Generator · Digit recognition · Fake sample · Real sample

P. Annan Naidu (✉) · S. Papa Rao
CSE Department, Aditya Institute of Technology and Management, Tekkali, Andhra Pradesh, India
e-mail: panaidu1411.cse@adityatekkali.edu.in

S. Rao Pappu
CSE Department, GITAM School of Technology, GITAM University, Visakhapatnam, Andhra Pradesh, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_58

1 Introduction

Computer vision is a new kind of computer field, through capturing the vision to make information expressed intuitively. Computer vision states the capability of a machine to understand images and videos. It mimics the capability of human vision by attaining, processing and analysing real-world data and synthesizing them into useful information. It uses a camera to capture images and videos to analyse, which can then be purposed for object recognition, motion estimation and video tracking. Computer vision is widely used in many fields such as industrial detection, digital image retrieval, biological image analysis and vehicle navigation [1], Internet of Things, Industrial Internet of Things and brain human interfaces [2]. Computer vision can perform the following tasks such as object identification, object classification and object tracking. Object classification used for classifying the objects based on visual content to classify on image or video to define object category; Object identification identifies particular object on image or video content; object tracking, to track object movements by processing the video content. In past and present digital era, machine learning used in different research fields such as health, education, medical diagnosis, medical image processing, bank and finance sector with its various models. Machine learning can perform tasks without explicit instructions. Machine learning can process the data and produce the results in terms of classification and prediction. There are number of machine learning models which are available such as supervised learning, unsupervised learning and semi-supervised learning models. Machine learning solution improvements around the data collection, data pre-processing, understanding data in form of training and test data, fit that train data into machine learning models, evaluate the model performance based on the test data to decide the model performance in form of accuracy. Machine learning and computer vision evolve in computing filed to bring the computers in the human competences for data detecting, data understanding and action activity based on the past and present outcomes. The basic work flow process of machine learning includes data collection, data pre-processing obtaining data features, understanding features, fit data into machine learning model, process the data for classification or prediction and model performance evaluation. In other end, computer vision process imitates the human capabilities, computer vision can capture images or videos through device called camera, and it also analysing data insights through computing devices and produce the meaningful interpretations. Machine learning and real time applications often used together effectively acquire analyze, and interpret captured visual data. Supervised learning algorithms deals with labelled data by using classification algorithms. The main objective of supervised learning is to understand data through training data set which includes both input and expected output attributes then test the accuracy with unseen dataset. In this algorithm, learning continues till it achieves an adequate level of performance.

Supervised learning algorithms are two categories such as classification and regression. The key aspect of classification is to classify the given data into different classes. For example, consider a student marks data, when we apply any classifier on this data, then the expected result is pass class or fail class. Classification models include support vector machine, decision tree, random forest and Naïve Bayes, whereas regression models are linear regression and logistic regression. Regression is used to define the relationship between independent variable and dependent variable to predict the dependent variable also known as target variable. Regression also identifies the relationship between the significant impact of single and multiple independent variables with dependent variable. The main task of supervised learning is classification and predicts a target numeric value. The main applications of supervised learning algorithms are speech recognition, spam detection, object recognition for vision, face detection, image classification and text categorization.

Unsupervised learning is a machine learning approach; it is used to detect the unseen patterns in the given data set without known label data. Supervise learning models deal with labelled data and it mainly deals with unlabelled data. Clustering is one of the important concepts, clustering is a process of grouping unlabelled data, in which the unlabelled data can be divided into different groups with similar data and each group is called a cluster. For example, consider a student data, in which data can be divided into different clusters and each cluster represents similar data points like department-wise students, admission-wise data, etc. Clustering methods are classified into two types such as hard clustering and soft clustering. Hard clustering means data points belong to only one cluster, soft clustering means data points belong to any other clusters. There are different types of clustering models but this paper mainly focused on K-Means clustering; K-Means clustering is a partitioning clustering, in which data set can divided in form of non-hierarchical clusters and it is also known as data point centroid-based model. Clustering is an unsupervised machine learning tool used to determine different patterns of data set. Clustering is a way of segregating the data into different groups with similarity and these data groups are called as clusters. Clustering algorithms are used to classify data into specific groups. Grouping unlabelled data is called clustering. As clustering deals with unlabelled data, it relies on unsupervised learning. Clustering algorithms are used in various applications such as image segmentation, social data analysis, market analytics. Supervised learning approach is used for prediction and classification based on the given data set. Examples for prediction are like stock market prediction, weather forecasting, etc. whereas for classification, image classification, credit card detection and email spam detection. Clustering algorithms can automatically recognize the pattern inside the data so that data can be analysed without depending on label data. Unsupervised learning deals with unlabelled data which means there are no target or outcome variable and unknown relationship between input variable and outcome variable. The actual steps involved in machine learning process are as follows, such as data collection, data preparation, select a model, training a model, model evaluation, parameter tuning and prediction.

2 GAN-Generative Adversarial Networks

Introduction section discussed briefly about machine learning process in order to make clear understanding and also which is relevant to the GAN model [3]. Machine learning process involves the train data, fit data into model and evaluates the model performance by mapping input labels with the class labels is known as classification based on supervised learning approach, and also traditionally referred as discriminate model, whereas in unsupervised approach, encapsulate the distribution of input variables may be able to be used to create or generate new examples in the input distribution, is referred as generative model. So in machine learning process, discriminate model and generative model take place at two different approaches. GAN is a machine learning community framework introduced by Good Fellow and his teammates [3], GAN is a generative model, in which neural networks are trained to compete each other. GAN model learns from training data set as it is with same features of training data set. GAN model process flow architecture consists of two neural networks; generator and discriminator and these two networks are in like game scenario. That is generator is used to produce replica data from the actual data set and discriminator is looking for to differentiate the replica data and the actual data. Sometimes may be the generator make fools to discriminator by creating better replica data, and at same time, discriminator also gets a better distinction among replica and actual data. In order to minimize the loss and gain among these two networks in game scenario, GANs used probability distribution to find the distance between and replica data and actual data. In GAN model, generator works like unsupervised learning and discriminate model works like supervised learning. GAN Lab [4] was introduced by Kahng and his team. GNA Lab describes the complete work flow of GANs functionalities. Figure 1 depicts the commonly used GANs architecture.

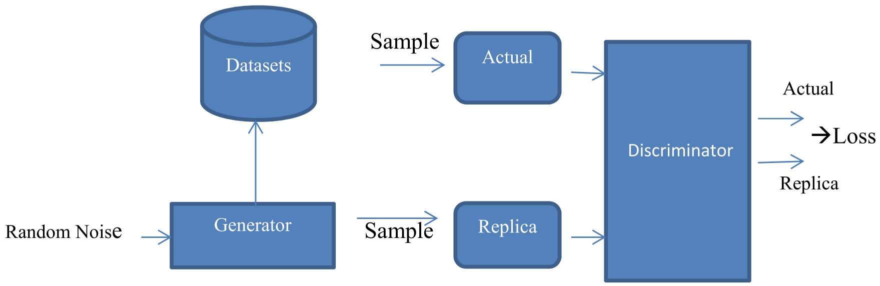


Fig. 1 GAN architecture

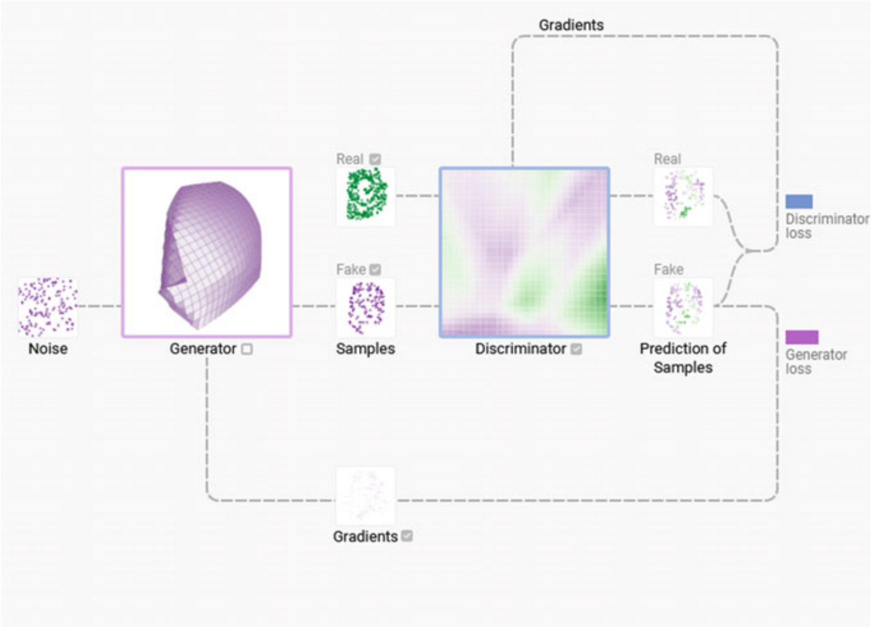


Fig. 2 GAN model view [4]

Figure 2 represents the GAN model work flow through GAN Lab, which starts with trained data, generator produces data in terms of fake data from real data samples and other end discriminator is used to predict the real and fake data from the generated data. In GAN Lab, select random data from the samples or create a sample data and provided it to GAN model to differentiate the fake and real data and the process represents in Fig. 3, which depicts the real samples and fake samples and metrics represents the function loss of discriminator and generator.

GAN model applied on thousands of handwritten digits, collected from MNIST data set [5]. The main objective is to train GAN model with the above said data set to generate replica data using Tensorflow-GAN estimators (TF-GE). By using TN-GE, loaded data set from MNIST database through Python library support.

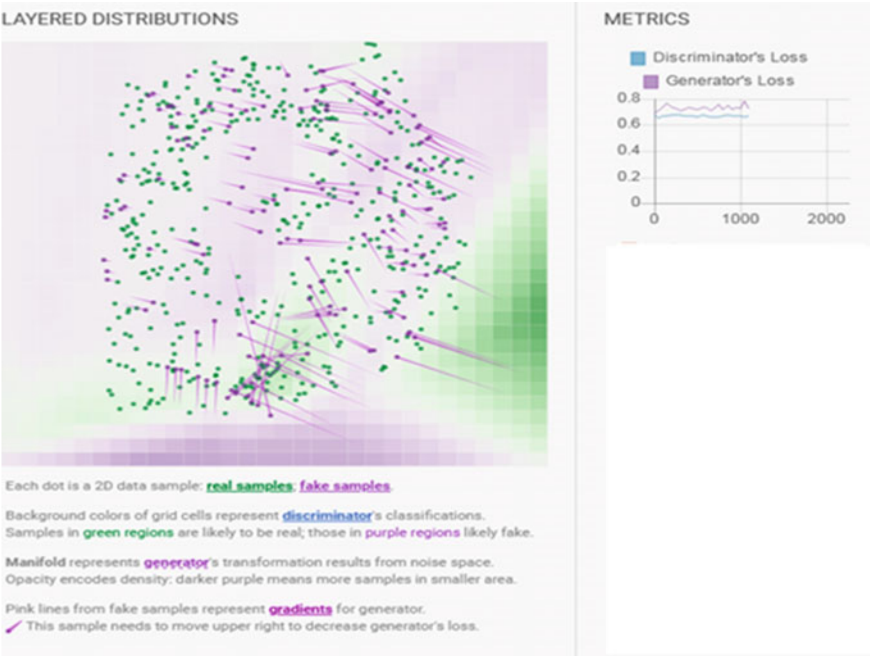


Fig. 3 Layered distributions and metrics [4]

3 Results

The proposed work includes the following steps:

1. Install TF-GE-related Python libraries.
2. Load data set from MNIST database through Tensorflow_data sets in Python.
3. Train GAN model by using Tensorflow-GAN estimators (TF-GE).
4. Create two neural network architectures such generator and discriminator.
5. Evaluating GAN model with evaluation metrics such as inception score [6] and Frechet inception distance, here inception score is used to measure the quality of generated images and Frechet distance is used to measure closeness of the generated images with real images.
6. Repeatedly generated data and the score will be displayed in Fig. 5.

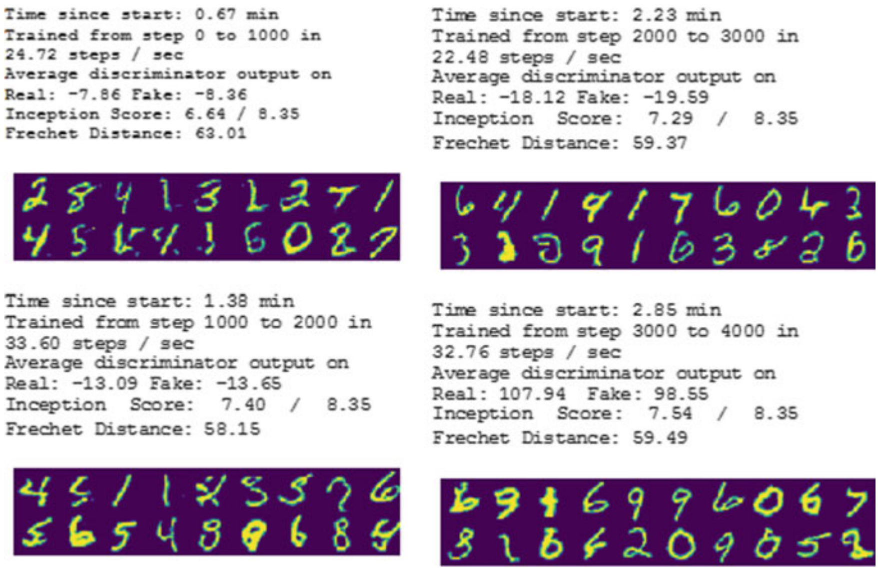


Fig. 4 GAN-based handwritten digit recognition

4 Conclusions

This paper aims to provide deep understanding of machine learning and computer vision approaches and its applications with its advanced research. Studied existing research work papers on machine learning and computer vision and the same will be presented in the above sections. Apart from this, this paper also discussed as an application part, that is the importance of Generative Adversarial Network (GAN), GAN Lab implementation. Implemented GAN model with handwritten digit data and results shown in Figs. 4 and 5 respectively. As per result concern, GAN model executed first iteration with data samples and produced results along with GAN model performance model metrics and the same will be available in Fig. 4, then GAN model repeatedly executed for three iterations and generated scores, represented in Fig. 5. Finally, the created GAN model can be used for different images to train GAN model using TF-GE and evaluate model performance with inception score and Frechet distance.

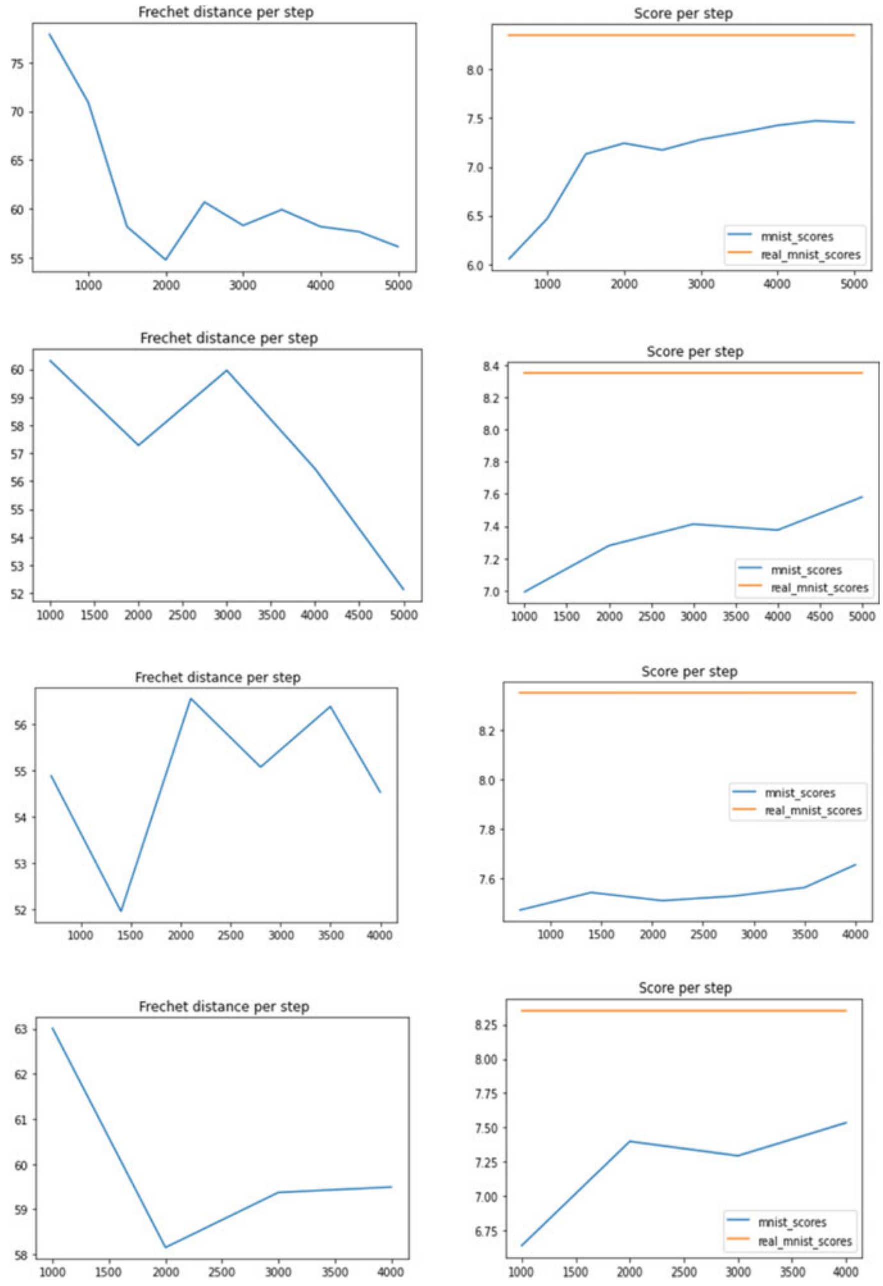


Fig. 5 Metrics for handwritten digit recognition system

References

1. Ci Y, Jiao F, Tu W, Fang J (2021) Analysis on development tendency of computer vision and graphics based on bibliometrics. J Phys: Conf Ser 1883:012054. https://doi.org/10.1088/1742-6596/1883/1/012054
2. Khan AI, Al-Habsi S (2020) Machine learning in computer vision. In: International conference on computational intelligence and data science (ICCIDS 2019). Procedia Computer Science, vol 167, pp 1444–1451
3. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial networks, 10 June 2014. arXiv:1406.2661v1 [stat.ML]
4. Kahng M, Thorat N, Chau P, Viégas F, Wattenberg M (2019) GAN lab: understanding complex deep generative models using interactive visual experimentation. IEEE Transactions on Visualization and Computer Graphics 25(1) (VAST 2018)
5. https://en.wikipedia.org/wiki/MNIST_database
6. Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X (2016) Improved techniques for training GANs. Machine Learning (cs.LG), 10 June 2016. arXiv:1606.03498v1. https://doi.org/10.48550/arXiv.1606.03498

Pending Receipts RPA Bot

Soma Karthikeya and N. Meenakshi

Abstract The term "robotic process automation," or RPA for short, refers to a software system that may automate a variety of activities. Users of software may take advantage of RPA to build software robots, sometimes known as "bots," that can learn, imitate, and eventually carry out rules-based business processes. Users are able to create bots by modelling their behaviour after human digital activities thanks to RPA automation. After you have shown what you want your bots to do, you can step back and let them take over. Robotic process automation software bots are able to interact with applications or systems in the same manner that humans do; the main difference is that RPA bots can work continuously, nonstop, far quicker, and with a reliability and accuracy rate of one hundred per cent. The evaluation of RPA bots is presented in this document in its entirety. Automating invoice processing requires the usage of RPA bot, which is developed using Blue Prism.

1 Introduction

Pending Receipts: The main aim of the project is to reduce the human intervention as much as possible by automation of the process.

Since its very tough job for an individual to go through each and every invoice and checking the good receipt for the same then validating if everything in the documents are matching such as currency and amount. It is also difficult to manage the labour cost, human errors, and time taken. This is where automation comes into picture where we completely automate the procedure and complete all of the procedure in just few clicks.

S. Karthikeya (✉) · N. Meenakshi
Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur, Chennai, India
e-mail: sk7515@srmist.edu.in

N. Meenakshi
e-mail: meenaksn@srmist.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_59

Pending Receipts process is a 3-way match between invoices in Coupa, GR in SAP, and the invoice attachment within Coupa. A Coupa BOT is in place to post these invoices by using Rossum for OCR. Categorization of invoices happens based on the reports extracted from Coupa and SAP, followed by actions that are being taken such as posting of invoices for the perfect matches, escalation matrix in case of data mismatch.

1.1 Problem Statement

To automate the process of payment validations (IR GR PO) and reduce as much as human intervention as possible, labour cost, errors, and time.

According to a survey, businesses are able to do better with the rise of automation: Despite the fact that 83% of executives said their organizations presently invest in and/or utilize automation or AI technologies, 78% are very or somewhat inclined to spend more in automation to mitigate the effects of the labour crisis.

In 2022, about seven out of ten CEOs aim to expand their investment in automation technologies compared to 2021.

2 Literature Survey

Your company will have more time to concentrate on its key goals as a result of automation's ability to cut down on time, effort, and costs while also minimizing the number of mistakes caused by human processes. It is possible to finish jobs that are repetitive more quickly. When operations are automated, the outcomes are guaranteed to be of a high quality since there is no room for human mistake in the performance of each activity.

Surendhiran Tamilalagan Automation minimizes time, effort, and expense while decreasing human mistakes, allowing your organization more time to concentrate on its key goals. Repetitive chores may be performed more quickly. The elimination of human error via the automation of processes promotes high-quality outcomes, since each activity is executed similarly.

Sagar Sahu, Sania Salwekar, Atharva Pandit, Manoj Patil a recent endeavour to construct an automobile application to revamp the billing process in finance operations is presented in this article. As a prime illustration of the technology's ability to increase efficiency, Robotic Process Automation (RPA) is often used to a number of financial and accounting procedures, including the invoicing process. Once invoices are processed, RPA Data Bot will automate data entry, error reconciliation, and some of the decisions required by finance personnel. Simultaneously, automation is able to decrease failures in these operations and reduce the need for human exception management. UiPath's RPA Data Bot is able to continuously monitor a folder, in which PDF invoices are stored by employees (or another Data Bot). Once robots detect the existence of an invoice inside a folder, they begin extracting data from the document. Using clever Optical Character Recognition (OCR) and language processing skills, Data Bot can extract the information shown on an invoice.

Chac´on-Montero, A. Jim´enez-Ram´ırez, J.G. Enr´ıquez [1] In the article, the authors have implemented RPA in various processes and its results are discussed. It was observed that RPA in these processes is 78% more effective than humans doing the same job. The RPA bot is also tested. A virtual environment is created, and the RPA bot is tested in this virtual environment. Creating a virtual testing environment saves money and other resources.

Nataliya Yatskiv, Solomiya Yatskiv, Anatoliy Vasylyk [2] suggested the use of AI with RPA for software testing. Using AI increases productivity and saves resources. Computer vision is used for recognizing objects and collecting its data. RPA tools are divided depending on their functionality and AI support level. The major tools are included in: Excel automation and macros, programmable solution bots, self-learning tools, and cognitive automation tools.

Enríquez et al. In, it is concluded that interest in RPA technology is growing rapidly and that several market solutions are currently available for RPA. There are several digital libraries established for RPA stages, except for the analysis step owing to a lack of industry interest.

Martins et al. Automating the detection and classification of application software interfaces in real time using convolutional neural networks was suggested.

Ying Li, Muthu Muthiah, Arindam Routh, Chitra Dorai explain their intention to build a cognitive application to alter the Procure to Pay (P2P) system by delivering domain-specific solutions. The Procure to Pay (P2P) process comprises two primary tasks, namely the transmission of transactional data to the provider and the management of data that contains information about orders and payment for a product or service. A three-month lengthy use monitoring revealed good results. Since deploying this application to a client's production environment, the team's productivity has increased substantially.

Enes Aslan, Ethem Unver, Tugrul Karakaya, Yusuf Sinan Akgul In their work, they developed a novel technique for invoice parsing that removes invoice classes and consists of a two-phase optimization structure. The first step employs individual invoice component detectors, such as SVM, maximum entropy, and HOG, to generate candidates for the various sections of various invoice types. At the second stage, the objective is to separate a bill into its component components and then organize them. As PBM is built on optimization, it can manage any form of billing. The suggested approach is evaluated using actual bills and is determined to be promising for use in the real world.

Xufeng ling, ming gao, dong wang. Processing of documents using artificial intelligence based on robotic process automation and machine learning RPA refers to a system that can carry out activities that are static and repetitive. It comes in very helpful for reading PDF files, removing data from such files, reading data from images, and other similar tasks. RPA makes it easier to collaborate with a wide variety of technologies, including AI, machine learning, natural language processing, and others. In commercial settings, it serves the same purpose as an employee that never clocks out and is always available. During this inquiry, RPA and AI will be used to test out intelligent document processing as well as document reading. Using computer vision, the page number and document's title are first extracted. Next, natural language processing is utilized to extract textual characteristics and then mapped based on cosine similarity. A machine learning method is then utilized to generate a summary of the document. There is also a human-interaction component, where the person may check if the description extracted from the document is correct or not.

Ghosh Automation using Robotic Process Automation (RPA) Robotic process automation provides a multitude of benefits. The capacity to record and replay mouse and keyboard motions is one of its most significant capabilities. This article explains the software bot designed to monitor employee email throughout the on-boarding process. The software bot monitors the mailbox, and when a new message arrives, it alerts the user.

When an email is received, the system will get the appropriate data from it. This data may include the employee's name, the date they started working, their assets, and other information. The information is then saved in the database, and an alert is sent to the appropriate division on the new employee's on-boarding. The whole process of reading the email, obtaining the data from the database, and sending a notice took just 43 s when the bot was deployed, but it would have taken over five minutes to accomplish the operation manually. Even if there were hundreds of emails, the bot would handle them all quite efficiently and do this duty.

Karthick, K.B. Ravindrakumar, R. Francis, S. Ilankannan Optical character recognition is one of the automated identification systems that fits the automation needs of different applications (OCR). With OCR, a computer can interpret information in any format from natural landscapes or other resources. Character identification while typing and printing is straightforward due to their well-defined size and form. Individuals' handwriting differs in the above ways.

T.A. Bayer, H.U. Mogg-Schneider, Daimler-Benz T. presented a generic system for processing bills that automatically retrieved the needed components from any form layout in any domain. The provided system has two components: an OCR tool and a FRESCO repository. FRESCO has a lower mistake rate, hence it is used. FRESCO component includes domain-specific information. This generic system is used for billing connected to health insurance.

Institute of Exact Sciences and Informatics Discusses the current experimental results on the proposed method for recommending and aggregating related content utilizing the SCORM-defined relation metadata category. This method makes use of ontologies, automated annotation of metadata, retrieval of information, and text mining. It is a prototype of a computer system that applies the recommended methodology to a selection of learning objects and provides results to evaluate the usefulness of the methodology. The findings point to the fact that the proposed approach is both practicable and successful in its pursuit of the goals that are wanted.

Shan Feng, Zhu Li, Yiling Xu and Jun Sun We have created a one-of-a-kind hashing technique that is based on the combination of Fisher Vector aggregation and deep convolutional network feature mappings. Two of the most important contributions made by this work are the aggregation of Fisher Vectors based on the characteristics of an underlying Gaussian Mixture Model (GMM), as well as the creation of a hash via the direct binarization of Fisher Vectors and component optimization.

Munidar P. Singh The enterprise-focused conceptual framework that was suggested before. Abstractions of goals, commitments, and methods that have been well examined are used by Chabot. In contrast to traditional IFTTT frameworks, which make the maintenance of Chatbots prohibitively resource- and cost-intensive, the use of these abstractions enables the building of complex Chatbots in a flexible manner. According to the findings of the study that are provided in this article, an engineering Chabot is made up of a knowledge base, a conversation manager, an inference engine, and a planner.

A study titled "Content Aggregation by Platforms" was recently published. Uses—well-researched abstractions of objectives, commitments, and plans to provide a conceptual framework for enterprise Chabot. Instead of utilizing the current IFTTT frameworks, which make maintenance prohibitively resource and expense heavy, the complicated Chabot may be constructed in a flexible way with the help of these abstractions. An engineering Chabot is made up of the following elements, according to the research presented in this paper: the dialogue manager, dialogue management, inference engine, knowledge base, and planner.

Dan He and Douglass S. Parker talk about a model for tailored content aggregation that takes both user access rate and the pace at which new material is posted into account.

Additionally, this model takes into account the delay period for these missing posts in addition to the quantity of missing posts. To reduce the anticipated aggregate delay time across many data sources, they created an efficient resource allocation technique as well as an optimal retrieval scheduling system.

3 Methodology

(1) Blue Prism

Blue Prism here is used to automate the process which acts as an RPA bot which efficiently performs excel automation and web automation and SAP automation.

(2) Rossum

Rossum acts an OCR tool which basically identifies the text or important fields from the invoice receipt. To perform 3-way match.

(3) Sap

SAP basically stores all the purchase document number along with invoice details. This data which is respective T-codes are used to extract the data.

(4) Pre-processing

The data is being pre-processed and cleaned up accordingly in order to pass the data to the next steps.

(5) 3-Way Match

Matching the data from OCR text via Rossum (Invoice Receipt) and SAP data along with the good receipt data from Coupa.

(6) Posting of Invoices

If there is a perfect match, the posting of invoices is done using selenium via Coupa else these invoices are categorizes into necessary categories and escalation mails are sent accordingly.

4 Activity Flow and Architecture Diagram

See Figs. 1 and 2.

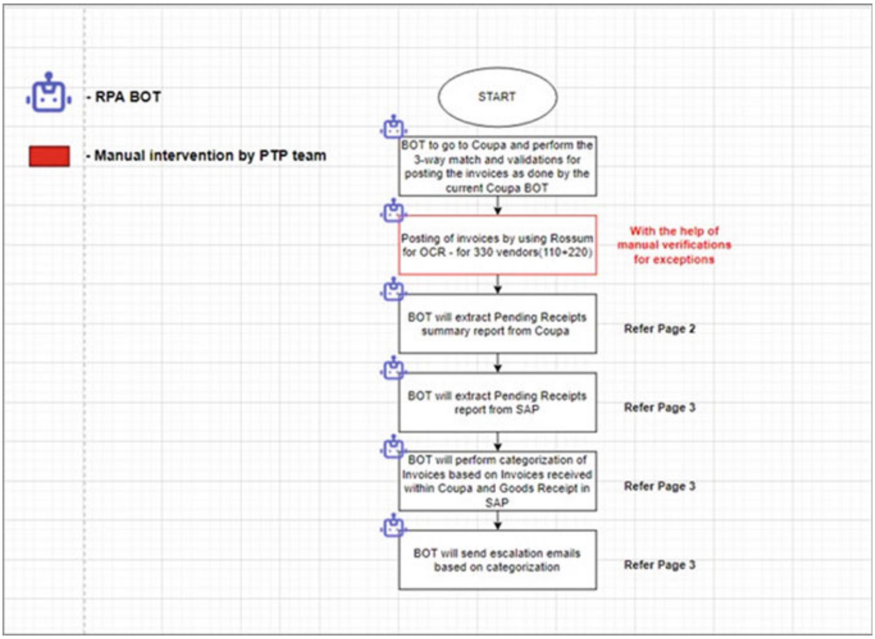


Fig. 1 Blue Prism flow diagram

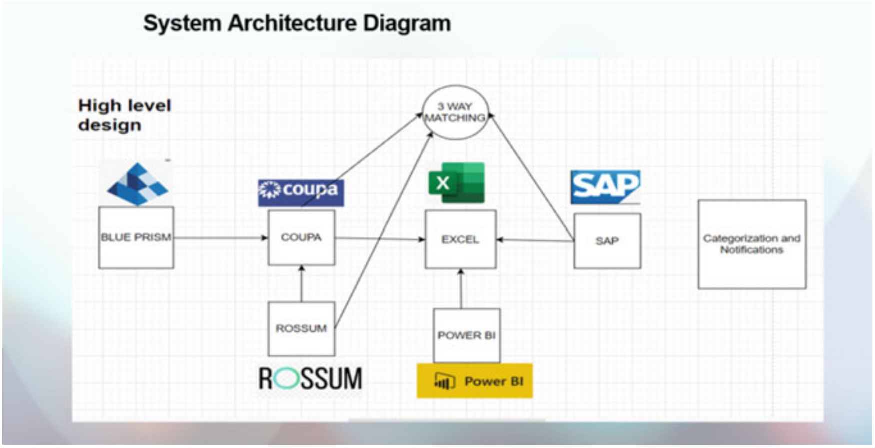


Fig. 2 System architecture diagram

5 Results

The given project releases one FTE as this automation reduces about 90% of the work load (Figs. 3 and 4).

RPA Tools

Automation Anywhere: Automation Anywhere is a well-known RPA vendor that provides robust and easy-to-use capabilities for automating any complex business

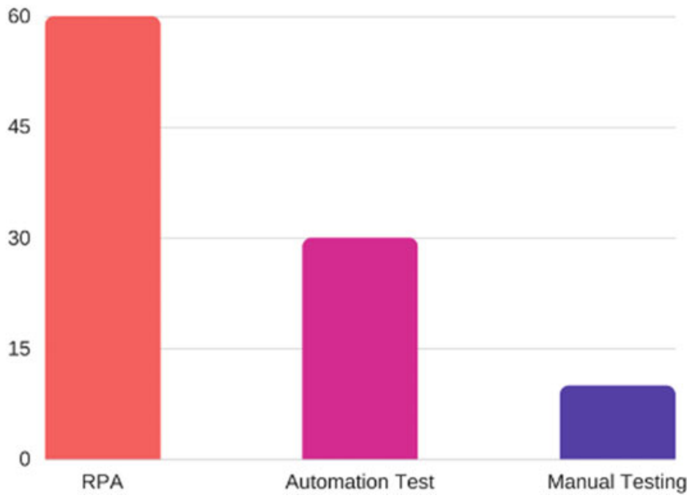


Fig. 3 Invoice processing with automation is studied

Time Saved	Frequency of Task					
	50/Day	5/Day	Daily	Weekly	Monthly	Yearly
1 Second	4	0	0	0	0	0
5 Seconds	18	2	0	0	0	0
30 Seconds	108	11	2	0	0	0
1 Minute	217	22	4	1	0	0
5 Minutes	1,083	108	22	4	1	0
30 Minutes		650	130	26	6	1
1 Hour		1,300	260	52	12	1
6 Hours				313	73	6
1 Day					292	24

Break even time in hours per year that the process will run.

Fig. 4 Bot run timings are being shown

process. This utility combines all of the essential features. It blends RPA with cognitive features such as language comprehension and unstructured data reading. It is a web-based management system that allows businesses to govern and manage end-to-end automated business processes. It enables the automation of a wide range of tasks, ranging from simple Windows settings to advanced networking, and remote database processes.

Blue Prism: Blue Prism is a robotic process automation (RPA) solution that gives companies access to a virtual workforce. It makes it possible for businesses to automate manual, rule-based, and repetitive business processes rapidly and at a cost that is reasonable. It enables the automation of tasks via the use of a drag-and-drop interface. Blue Prism is built on the.NET Framework developed by Microsoft. It is able to automate any programme and is compatible with any platform (mainframe, Windows, WPF, Java, web, etc.) provided in an extremely diverse range of methods that are applicable (terminal aper, thick shopper, skinny shopper, browser, Citrix and net services). It was developed with a multi-environment readying model in mind (development, testing, staging, and production), and it has both physical and logical access restrictions built into its architecture. A centralized task management interface and method amendment distribution architecture are both included in the Blue Prism RPA software system. These features combine to provide high levels of visibility and administration. The company receives additional management support in the form of a centralized model for the creation and reuse of methods. Blue prism keeps a record of every system login, change in management action, as well as the selections and actions carried out by the robots in order to gather data, and conduct periodical operational analyses.

6 Conclusion

This article contains a briefing about what is RPA? What are its components and how to use and implement them for better outcome. In this paper, we have discussed the struggles and difficulties in manually processing invoices referring multiple sources. We have highlighted the various RPA tools available and their applications. Since Automation Anywhere satisfies our requirement, we chose it. Later, we have a detailed description about the various methods of extracting data from the invoices and the process of the working bot. Robotic Process Automation (RPA) is the optimal solution for replacing repetitive human actions and activities. This paper finally tells how Blue Prism can be used to replace manual invoice processing and mailing using PDF extract field and excel advance. By doing we can increase the accuracy and efficiency by a large extent and save ample amount of time.

References

1. Montero JC, Ramirez AJ, Enríquez JG (2019) Towards a method for automated testing in robotic process automation projects. In: 2019 IEEE/ACM 14th international workshop on automation of software test (AST), IEEE
2. Yatskiv N, Yatskiv S, Vasylyk A (2020) Method of robotic process automation in software testing using artificial intelligence. 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), IEEE

Abstractive Text Summarizer Using Neural Networks Algorithms

Dadi Ramesh, Kothandaraman D, Ramesh Chegoni, Sallauddin Mohmmad, and Syed Nawaz Pasha

Abstract Nowadays, the computation for various types of exams is increasing due to the increase in the population, so a person needs to study very hard and for a long time to pass or get a good rank in the exams. So, they need to study so much data and books, and to make study strain less the application "Text Summarizer" is used to convert huge amounts of data into small lines of data—the summarized data. This summarizer data gives the total meaning for the huge amount of data. This application is also used in the preparation of exams overnight. In overnight preparation, we need to prepare a lot of data in a small amount of time. So, using this, we can read and understand a lot of data in a small amount of time. We used seq2seq, LSTM modeling, and used attention machine. In our project, we got an accuracy of 82.54% for the 4515 records that we have taken.

Keywords Summarization · Seq2seq · Long short-term memory · Neural network

1 Introduction

Text summarization history is the one where there is always a fight for its development. The origins of it start from 1940 at the time of World War 2 to convert Russian language into English though development in this area has been in code and syntax. The main hype text summarization [1, 2] was got in 1990s. In those days, they used rule-based algorithms which used to say the importance and ranking of text so it is developed, and in 2016 they developed a model using seq2seq [3–5]. We used this model and another model in this proposed model.

D. Ramesh · Kothandaraman D (✉) · S. Mohmmad · S. N. Pasha
School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India
e-mail: dramanko@gmail.com

R. Chegoni
CMR Technical Campus, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_60

The main aim for text summarization is to get the most precise and useful information from a large document and eliminate the irrelevant and less important ones. Since there are many sites and blogs and news articles on the Internet, there is so much data to understand a topic, we need to read whole blog or site, it contains much of unnecessary data and unstructured data or a lot data that is relevant but you need to only understand the main idea of the data, so here the text summarization [6, 7] comes to play and it summarizes the entered data and makes a small text of the whole amount of data. Nowadays, the computation for various types of exams is increasing due to the increase in the population, so a person needs to study very hard and for long time to pass or get a good rank in the exams. So, they need to study so much data and books, and to make study strain less the application "Text Summarizer" is used to convert huge amount data into small lines of data the summarizer data [8–10]. This summarizer data gives the total meaning of the huge amount of data. This application is also used in preparation of exams overnight. In overnight preparation, we need to prepare a lot data in a small amount of time.

The major challenges with text summarization is to convert given text into vector that includes content. We have various text embedding methods in NLP like bag of words and TF-IDF [6, 11], Word2vec [3], sentence encoder, and BERT. In this, first two are statistical-based text embedding methods which will fail to capture semantics from text. And word2vec convert the text into vectors word by word and capture the semantics, but for words which meaning depends on its adjacent words it will give same vector. So in our approach, we used USE for text embedding.

2 Dataset

The dataset we used is Kaggle article summarization from https://github.com/sunnysai12345/News_Summary which consists of 4515 samples including article and short text in English language as given in Table 1. First, we preprocessed the dataset by removing special symbols, and we tokenized all the samples into sentences. Both article and short text are embedded to vectors using sentence encoder [11–13] of size 512 for each sentence. First, we tokenized all articles and short text to sentences and for that we used NLTK natural language processing library, which equalized all the articles by padding [14, 15]. Through dictionary module, we corrected all word spellings and removed special symbols from the both article and short text [16, 17].

Table 1 Sample vector of short text vector and dimension of vector

Dataset	Text vector	Dimension
News_ summary	[[0.01377844 − 0.09247074 0.01014971 … − 0.01349053 − 0.04146808 0.05626552] [−0.01370333 − 0.0240192 − 0.03880018 … − 0.05234249 − 0.06115109 0.05296136] [−0.05529318 − 0.02587196 − 0.00212097 … 0.02701825 0.02506788 0.00300164] … [0. 0. 0. … 0. 0. 0.] [0. 0. 0. … 0. 0. 0.] [0. 0. 0. … 0. 0. 0.]]	Sentence length * 512

3 Methodology

We have used the dataset of "Amazon fine food reviews," this dataset is legitimate and has frequently updating, since this developed dataset there in no missing values in the dataset, we have taken. All the boundaries and columns are good. If there are any missing values, these are the techniques that we will use:

- Remove the entire row (If missing values are less in number),
- Replace the missing value with either mean or median,
- Replace the missing value with most frequent value in the column (This is generally used only for large dataset).

4 Outliers Detection and Treatment

Anomalies are information focuses that do not fit the example of rest of the numbers. They are the very high or very low qualities in the informational index. A straightforward method for finding an exception is to look at the numbers in the informational index. We can likewise identify anomalies by Z-score technique, and its equation is given by:

$$Z = (x - \mu)/\sigma \tag{1}$$

where

- x is each value in dataset,
- μ is mean,
- σ is standard deviation.

5 Normalization

Normalization is a technique for organizing data in a database. Data normalization is the technique of rescaling one or greater attributes to the variety of zero to 1. In this method, the most important fee for every characteristic is 1 and the smallest fee is zero. It is vital as a database shown in Fig. 1.

$$x_{\text{nor}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

$$y_i = y_{\text{nor}}(y_{\max} - y_{\min}) + y_{\min} \tag{3}$$

Table 2 illustrates the hypermeters trained for LSTM proposed model. We used sentence-level embedding for text to vector conversion, and trained it for 50 epochs.

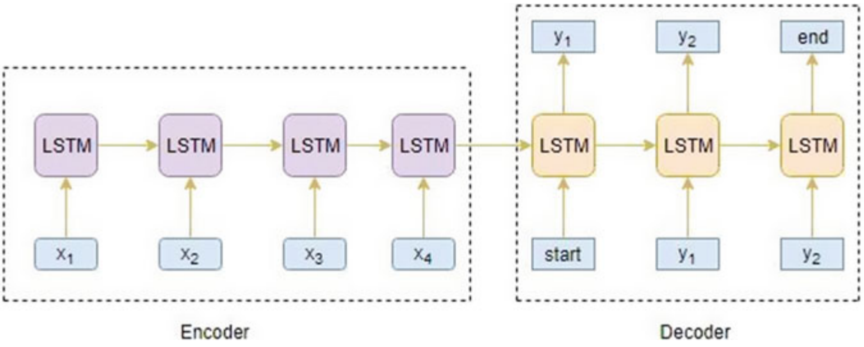


Fig. 1 Seq to Seq LSTM model

Table 2 Hyperparameters tuned for LSTM

Layer	Parameter	Value
Embedding	Sentence embedding	512 size vectors for each sentence
LSTM layers	No of layers	5
LSTM units	LSTM units	300
Hidden	Hidden units	200, 100
Dropout	Dropout rate Recurrent dropout	0.4 0.5
Others	Epochs Batch size Learning rate	35 32 0.001
	Optimizer	Adam

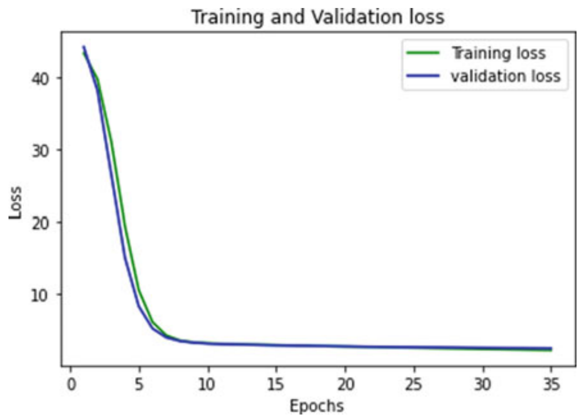


Fig. 2 Training and validation loss of LSTM model

6 Results Analysis

We trained the LSTM model with fivefold cross-validation, before that we separated all data into training and testing and validation in the ratio of 70, 20, 10. We got an accuracy of 82.54% for our model although it can be improved by taking the complete dataset which is five times of the present dataset. The training and validation loss of proposed model is shown Fig. 2. From Fig. 2, it is clearly observed that our model is not over fitted and under fitted.

7 Conclusion

Text summarization is generating short text from given article or text. In this paper, we proposed different approach that is sentence-based embedding and training recurrent neural network model with that achieved good result. In natural language type of applications, the hard task is to convert the given text into vectors, and we used universal sentence encoder for embedding text and converted it to vectors and trained it on LSTM, and achieved good results when compared to other models. In future, we will try to improve the accuracy as well we test our models on adversarial text to test the model.

References

1. Abualigah L, Bashabsheh MQ, Alabool H, Shehab M (2020) Text summarization: a brief review. In: Recent advances in NLP: the case of Arabic language, pp 1–15
2. Kumar AKSHI, Sharma ADITI (2019) Systematic literature review of fuzzy logic-based text summarization. Iranian Journal of Fuzzy Systems 16(5):45–59

3. Widyassari AP, Rustad S, Shidik GF, Noersasongko E, Syukur A, Affandy A, Setiadi DRIM (2022) Review of automatic text summarization techniques & methods, Journal of King Saud University—Computer and Information Sciences 34(4)
4. Sunitha C, Jaya A, Ganesh A (2016) A study on abstractive summarization techniques in Indian languages. Procedia Computer Science 87
5. Kiyoumarsi F (2015) Evaluation of automatic text summarizations based on human summaries. Procedia—Social and Behavioral Sciences 192:83–91
6. Afzal M, Alam F, Malik K, Malik G (2020) Clinical context–aware biomedical text summarization using deep neural network: model development and validation. J Med Internet Res 22(10)
7. John A, Premjith P, Wilscy M (2017) Extractive multidocument summarization using population-based multicriteria optimization. Expert Syst Appl 86:385–397
8. Elhadad N (2004) User-sensitive text summarization thesis summary. American Association for Artificial Intelligence (www.aaai.org), USA
9. Kiyomarsi F, Esfahani FR (2011) Optimizing Persian text summarization based on fuzzy logic approach. In: 2011 international conference on intelligent building and management
10. Isfahani FR, Kyoomarsi F, Khosravi H, Eslami E, Tajodin A, Pooya KD (2008) Application of fuzzy logic in the improvement of text summarization. Iadis International Conference Informatics
11. Wei W, Ming Z, Nie L, Li G, Li J, Zhu F, Shang T, Luo C (2016) Exploring heterogeneous features for query-focused summarization of categorized community answers. Inf Sci 330:403–423
12. Tayal MA, Raghuwanshi MM, Malik LG (2017) Atssc: development of an approach based on soft computing for text summarization. Comput Speech Lang 41:214–235
13. Zeng Z, Yan X, Yu B, Zhou F, Xu G (2020) Khmer multi-document extractive summarization method based on hierarchical maximal marginal relevance. Journal of Hebei University of Science and Technology 41(6):508–517
14. Kothandaraman D (2017) Context-aware human tracking system for internet of things. Anna University Chennai
15. Kumar BV, Arabelli RR, Mothe R, Kothandaraman D, Naik KS (2020) Smart traffic control system for vehicles on roads using raspberry Pi. Journal of Mechanics of Continua and Mathematical Sciences 15(9):162–170
16. Mohmmad S, Dadi R, Kothandaraman D, Sudarshan E, Pasha SN, Shaik MA (2022) A survey machine learning based object detections in an image. AIP Conf Proc 2418(1):020024
17. Kothandaraman D, Praveena N, Varadarajkumar K, Rao BM, Dhabliya D, Satla S, Abera W (2022) Intelligent forecasting of air quality and pollution prediction using machine learning. Adsorption Science & Technology

Evaluation of Tongue Articulation in Children for Malayalam Vowels Utterance

Leena G. Pillai and D. Muhammad Noorul Mubarak

Abstract This work integrates automatic speech recognition (ASR) and articulatory phonetics within an artificial neural network (ANN) framework. The objective is to recognize and evaluate the place and manner of tongue articulation in children in the age group of five to ten. In the first phase, an ANN was designed with one hidden layer and ten neurons that classify the vowel sounds into the most suitable ten classes. The second phase follows the misclassification or error rate in the first phase and evaluates the tongue articulation using spectrogram formats. Systematic speech evaluation helps diagnose children's misarticulation and can train them to improve their speech quality. The feature set used for ANN classification includes Mel-frequency cepstral coefficients (MFCC) and formants patterns (F1, F2, and F3). The designed network has shown an average accuracy of 77%, and an evaluation study was conducted on different parameters in Malayalam articulatory phonetics (formants) of misclassified vowels.

Keywords Automatic speech recognition (ASR) · Articulatory phonetics · Formants frequency · Spectrogram · Neural network

1 Introduction

Vowel sounds are produced due to open airflow from the lungs and its path through the trachea and larynx. The tongue, lip, and jaw constriction control this open airflow and shape the sound to a differentiable vowel sound [1]. Vowel sounds are voiced because they are produced due to vocal fold vibration. In any language, vowels have stable and perceptible acoustic features [2]. Vowels are closely connected to the eccentric character of each language. The place and manner of tongue articulation and lip constriction determine the acoustic variability of each vowel. The tongue

L. G. Pillai (✉) · D. Muhammad Noorul Mubarak
University of Kerala, Trivandrum, Kerala, India
e-mail: leenabelieve@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_61

position and lip constriction determine each vowel's formants frequencies, affecting the quality of vowel sound [3].

The vocal tract acts as a resonator or a filter that shapes each speech sound. Different vocal tract structure forces to overwhelm some frequencies and, at the same time, underwhelms some frequencies. These resultant, overwhelmed resonance frequencies are called formants and are visible in the spectrogram [4]. The spectrogram is a three-dimensional representation of the compound frequencies of a sound. The x-axis represents time, the y-axis represents frequency, and the intensity represents relative darkness. The darker bands represent the peaks in the spectra (formants). In the case of vowels, each vowel sound can be classified with respect to frequency formants 1, formants 2, and formants 3 [5].

The speech recognition system for children is much more complicated than that of adults. The developing variability of the vocal tract system and the natural misarticulation characteristics of children lead to this complexity [6]. This work concentrates on the diagnosis of tongue articulation in children aged five to ten. Vowels are the preliminary studied sound in any Indian language because it requires only limited tongue articulation. Hence, Malayalam is the native language in Kerala; all target children are familiar with the Malayalam vowels.

1.1 Malayalam Vowel Classification

Vowels are classified according to the activities involved in the tongue portion, the position of the tongue against the palate, lip constriction, and duration of utterance. Figure 1 describes the Malayalam vowel classification.

The tongue height represents the vertical position of the tongue against the palate. The first formants (F1) are inversely related to tongue height [7]. The closed vowels (tongue-high vowels) are uttered by placing the tongue relatively close to the palate. Therefore, the F1 value will remain low compared to middle or low (open vowels) tongue-positioned vowels. Tongue backness defines the portion of the tongue actively involved in the utterance against the back of the mouth. The second formant (F2) is

Fig. 1 Malayalam vowel classification

Tongue Backness

Tongue Height	Duration	Front	Central	Back
High	Short	ഇ i		ഉ u
	Long	ഈ i:		ഊ u:
Mid	Short	എ e		ഒ o
	Long	ഏ e:		ഓ o:
Low	Short		അ a	
	Long		ആ a:	

related to the degree of tongue blackness [5]. The F2 frequency remains relatively high in the front vowels and relatively low in the back vowels, but it is usually affected by lip constriction (rounded). Lip constriction is easily visible. The third formant frequency (F3) is related to lip rounding, but the other formants, especially F2, highly influence the F3 frequency.

2 Methods

2.1 Dataset and Feature Extraction

Ten Malayalam monophthongs vowels were collected from 56 children in the age group of five to ten. The recorded sampling rate is 16 kHz. Two different ways of approaches are applied for recording and initially demanded them to utter the short as well as long vowels. Typically, children by heart the vowels in a rhythmic sequence and follow the same rhythm whenever they recollect. Therefore, another method is used to overcome this pattern. The vowels were displayed on the screen one by one, and asked the children to read. The children are distracted easily, so the speech sounds were collected in a relatively noiseless environment.

A feature vector consists of 16 acoustic features, 13 MFCC coefficients, and 3 formants frequencies (F1, F2, and F3). The Mel-frequency cepstral coefficient (MFCC) is the most commonly used acoustic feature for automatic speech recognition [8–10]. The 13 MFCC coefficients were extracted from 25 ms windows with an overlap of 10 ms.

2.2 Training and Testing

The neural network was designed with 16 input units and one hidden layer with ten neurons, and the output classification is 10 (10 Malayalam monophthongs vowels) (Fig. 2). Vowel utterances from 36 children were used for training, and 20 children's utterances were used for testing. This work is more concentrated on the error rate and intended to diagnose the cause of the error.

3 Result and Discussion

The trained model tested with 200 vowel utterances and has shown an accuracy of 76.7%. It has been noticed that the long and its corresponding short vowels have similarities in features. Five short vowels and their corresponding long vowels are

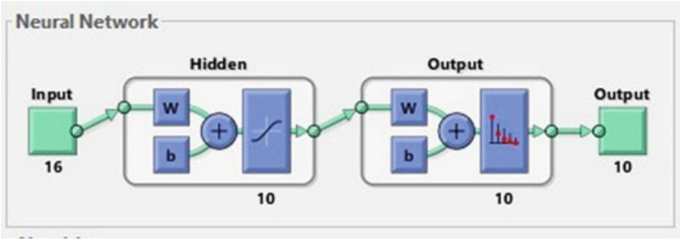


Fig. 2 Designed network model

considered. The only difference between the long and their corresponding short vowels is duration.

The long and short vowels share some common features because the articulation parameters of both long and short vowels are almost exact. Therefore, misclassification between the short and their corresponding long vowel is quite common, even in adults, because the significant difference is in duration. However, this duration is different from person to person. Hence, the misclassification between the long and short vowels in the same classes (e.g., a and a:) is reasonable. The place and manner of articulation required for vowels in a different class (e.g., a and i) are distinguishable, and their misclassification (e.g., a and i) is addressed in this work.

As described in Table1, the vowel classes low and central, mid and front, and high and front achieved an accuracy of 100%. Even though these classes have intraclass misclassification (e.g., a and a:), they do not have interclass misclassification (e.g., a and i).

The accuracy of high and back vowels is 69% and for the mid and back vowels is 87%. The 31% of high and back vowels are misclassified as mid and back vowels, and the 13% of mid and back vowels are misclassified as high and back vowels. The back vowels have shown the highest misclassification. The back vowels are uttered

Table 1 Interclass accuracy and misclassification matrix

Height and backness	High and front (i), (i:)	High and back (u), (u:)	Mid and front (e), (e:)	Mid and back (o), (o:)	Low and central (a), (a:)
High and front	100	0	0	0	0
High and back	0	69	0	31	0
Mid and front	0	0	100	0	0
Mid and back	0	13	0	87	0
Low and central	0	0	0	0	100

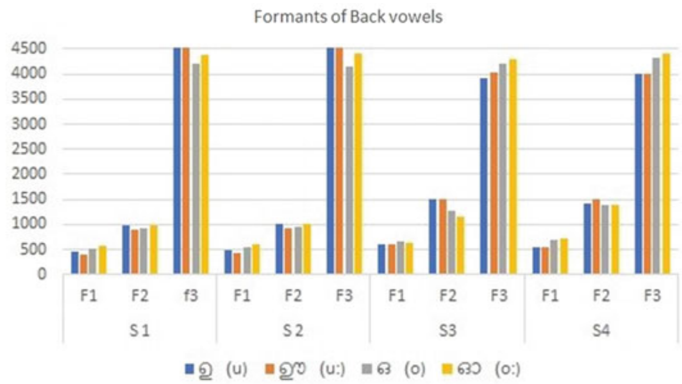


Fig. 3 Formants of misclassified vowel

by placing the tongue at the back of the mouth. In the case of high and back vowels (u and u:), the tongue should keep high and the position at the back, whereas in the mid and back vowels (o and o:), the required tongue height is mid, and the tongue needs to be placed at the middle of the mouth. Some children need assistance to take their tongues back correctly.

Figure 3 shows the formant frequencies of four speakers in the misclassified vow el. The S1 and S2 speakers are correctly classified speakers, and the S3 and S4 are the misclassified speakers. The F2 frequency of the back vowel remains relatively low. However, in S3 and S4, the F2 is around 1500 Hz and identified as they need to take their tongue at needed high. The manner of tongue height will affect the F1 and F3 also. The front–middle movement of the tongue is quite common. Therefore, the mid and front vowels are easy for children, but some children face difficulty in uttering back vowels.

4 Conclusion

This work focused on the evaluation of Malayalam vowel articulation of children in the age group of five to ten. The advancement in speech technology is occurring day by day. However, speech technology for children is at its infant level because the speech corpora of children are much more complex than that of adults. Hence, children are highly influenced by technology. The technological aid would be useful for children's speech and language development. In this work, the neural network classified 10 Malayalam monophthongs with 77% accuracy. The remaining 23% of error occurred in vowels considered based on their articulation parameter. The intra-class misclassification can be ignored as they require a similar articulatory parameter. The interclass misclassification occurred between the back vowels. Evaluating the F1, F2, and F3 frequencies, it is identified that the children face difficulty placing their tongues back during articulation. Articulation diagnosis is essential in speech and language development training. The tongue is the most versatile but not visible articulator during speech production. An effective tongue-tracking system would be a helpful tool for a speech-language pathologist in diagnosis and therapy.

References

1. Ionescu CM (2013) The human respiratory system: an analysis of the inter-play between anatomy, structure, breathing and fractal dynamics. Springer Science & Business Media
2. Johnson K, Ladefoged P, Lindau M (1993) Individual differences in vowel production. J Acoust Soc Am 94(2):701–714
3. Esfandiari N, Alinezhad B, Rafiei A (2015) Vowel classification and vowel space in Persian. Theory Pract Lang Stud 5(2):426
4. Zue V, Cole R (1979) Experiments on spectrogram reading. In: ICASSP'79 IEEE international conference on acoustics, speech, and signal processing, vol 4. IEEE, pp 116–119
5. Kent RD, Vorperian HK (2018) Static measurements of vowel formant frequencies and bandwidths: a review. J Commun Disord 74:74–97
6. Bhardwaj V, Ben Othman MT, Kukreja V, Belkhier Y, Bajaj M, Goud BS, Hamam H (2022) Automatic speech recognition (ASR) systems for children: a systematic literature review. Appl Sci 12(9):4419
7. Huber JE, Stathopoulos ET, Curione GM, Ash TA, Johnson K (1999) Formants of children, women, and men: the effects of vocal intensity variation. J Acoust Soc Am 106(3):1532–1542
8. Dhingra SD, Nijhawan G, Pandit P (2013) Isolated speech recognition using MFCC and DTW. Int J Adv Res Electr Electron Instrum Eng 2(8):4085–4092
9. Mistry DS, Kulkarni AV (2013) Overview: speech recognition technology, mel-frequency cepstral coefficients (MFCC), artificial neural network (ANN). Int J Eng Res Technol 2(10)
10. Mohamed FK, Lajish VL (2016) Nonlinear speech analysis and modeling for Malayalam vowel recognition. Procedia Comput Sci 93:676–682

E-Learning on Semantic Web

Ayesha Ameen and Ayesha Banu

Abstract E-Learning is the current trend that is gaining more significance post-pandemic, E-Learning has several distinctive features which has attracted many users across the globe, and there is a huge demand for E-Learning content. E-Learning is the delivery of learning material by any electronic gadget, and the communication channel used for the delivery of learning material is the World Wide Web. Currently, the content displayed on the World Wide Web is only understood by humans but not by machines, and in order to make machines understand the content and make intelligent decisions the World Wide Web can be replaced by the Semantic Web. In the current paper, SemELearn a Semantic Web-based E-Learning system is proposed that utilizes the underlying characteristic feature of the Semantic Web.

Keywords E-Learning · Semantic web · Ontologies · Inference

1 Introduction

E-Learning is an Internet-enabled learning, and it is the delivery of learning and training through digital resources like computers, tablets, or phones connected to Internet. E-Learning uses communication technologies to teach/train students/employee through various electronic media such as audio, video, multimedia, and visualization technologies [1]. E-Learning is otherwise known as virtual learning or Web-based learning. Instructors and students need not be available at the same place and time in E-Learning which makes it easy for learners to learn anytime, anywhere.

A. Ameen (✉)
Department of IT, Deccan College of Engineering and Technology, Hyderabad, Telangana 500001, India
e-mail: a.ameen@deccancollege.ac.in

A. Banu
Department of CSE (Data Science), Vaagdevi College of Engineering, Warangal, Telangana 506005, India
e-mail: ayesha_b@vaagdevi.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_62

Hence, learning and teaching become easier, simpler, and more effective by using various digital resources.

1.1 Growth of E-Learning

E-Learning is growing rapidly because for any queries people first search on the Internet rather than finding books or asking someone. According to Forbes, the size of global E-Learning market is expected to touch $325 billion by 2025. In today's context, what people understand for E-Learning is to get trained on any digital device and to gain knowledge by viewing an informative video, or by understanding an article published on Web or attempting a quiz all that is E-Learning.

With the growth in availability of digital content, everyone has accessed the E-Learning content to some or more extent, for knowing about new concepts and for getting acquainted to use modern technologies and gadgets. With so much information available on Web with in a microseconds, people want to explore the untraversed territories and want to be literates in all streams of knowledge. They do not want to be left behind at any cost in terms of technological advancement.

Web has supported E-Learning by enabling the information sharing and communication to be possible between instructors and students, and it has also helped the students to attend online lectures and access the study material provided by the instructor to be available in several multimedia formats. Web has also facilitated the use of several tools that can be used by instructors for creating study material, presentation, assignments, and quiz for students.

Hence, the Web has a crucial part in the growth of E-Learning, and if the Web is advanced to one more level in the form of the Semantic Web, then E-Learning can utilize the potential of Semantic Web technologies to provide more precise content to the students to suit their requirements, learning capabilities, and pedagogical needs. The term Semantic Web was devised by Sir Tim Berners-Lee for the Web of data which can be processed by machines, where the meaning of data is understandable by machines or machines are able to read the meaning of data rather than just representing the data.

1.2 Semantic Web

Semantics word in "Semantic Web" indicates meaning or understanding. Semantic Web differs from other technologies used in data representations such as databases or the World Wide Web which mainly emphasizes on the organization of data but not on the meaning of data. The underlying differences in Semantic Web technologies produce a diverse view on how displaying, storing, and querying of information is attained. Many applications that require huge amount of data to be processed from various sources can be benefitted from Semantic Web technologies. Three primarily standards are there in Semantic Web Resource Description Framework (RDF), Web Ontology Language (OWL), SPARQL Protocol and RDF Query Language (SPARQL). Reasoners also play a vital role in Semantic Web.

Semantic Web is built on W3C's Resource Description Framework. RDF is the formal language for explaining structured information. The foremost aim of RDF is to preserve the original meaning of data in the course of data exchange on the Web. RDF is used to describe the resources on the Web. Resources are represented in the form of triples. The relationship between subjects and objects is captured by the triples. Hence, RDF is a data model of the Semantic Web. Every piece of information in the Semantic Web is represented in the RDF form. RDF depicts the Web resources but does not describe the semantics of the resources. RDF Schema describes the semantics associated with Web resources. RDFS represents the web resources in the form of ontologies which consists of application defined classes and properties arranged in hierarchies. RDFS is not capable of representing complex relationships that exist among classes in ontologies, the usage of Web Ontology Language overcomes this drawback of RDFS.

OWL is more expressive than RDFS and supports the addition of more vocabulary for describing classes, properties, and relations between classes such as disjointness and cardinality constrains, characteristics of properties such as transitivity, symmetry. OWL is easy to use and understand, it is formally specified, i.e., the meaning of knowledge is defined precisely and can be used for automated reasoning.

SPARQL is RDF query language, it can be used for writing queries across varied data source. Queries in SPARQL comprises of triple patterns, disjunctions, conjunctions and optimal patterns. SPARQL can be implemented in several programming languages.

Recent advances in Semantic Web technologies have given rise to the need for sound and efficient methods for supporting reasoning on the knowledge scattered on Web. Ontologies support basic form of reasoning in as they are represented using OWL which is based on description logic. Further reasoning support can be provided in Semantic Web applications by incorporating logical formalism and corresponding inference engine. Inference based on free-form rules requires a rule engine. Rules are written for extracting the logical inferences which cannot be extracted using OWL DL reasoners.

2 Semantic Web for E-Learning

The main characteristics of E-Learning are delivery, access, responsiveness, symmetry, modality, authority, personalization, and adaptivity. These are summarized in Table 1 along with the how Semantic Web supports these characteristics features [2].

Table 1 Characteristic features of semantic web to support E-Learning

Characteristic feature of E-Learning	Brief description	Semantic web	Implemented in semantic web using
Delivery	Pull Agenda of learning is determined by students but not by the instructors	Items representing the knowledge are spread across the semantic web and are represented in the form of ontologies. Personalized course content can be created by extending, merging, and querying ontologies representing the domain	OWL, SPARQL
Access	Nonlinear Knowledge is accessed in any order that suits the current state of the students	Students can specify the current state of learning such as previous knowledge, goal of learning, and expected learning outcomes; they can also search the required material for learning by writing semantic queries. Ontological profiles of the students can also be created to gain more insight about the students. Ontologies can help in guiding students through various concepts	OWL, SPARQL
Responsiveness	Reactionary Responds to current problem/state	Semantic web has a formal language used by software agents for communication and dynamic delivery of course content. Each student has their own preferences, level of understanding, and requirements which are embodied in their personalization agent that interacts with the other agents	OWL, SPARQL
Symmetry	Symmetric Learning takes place as combined activity	Semantic web have the capability to serve as an integrated platform to all organizational business processes including the learning activities	OWL
Modality	Continuous Learning is a continuous activity and it occurs in a loop without any halts	Dynamic learning environment is created by active delivery of information by the personalization agent	OWL, SPARQL

(continued)

Table 1 (continued)

Characteristic feature of E-Learning	Brief description	Semantic web	Implemented in semantic web using
Authority	Distributed Course content are created based on the communication of the students and instructors	Semantic web is not centralized, and it is to a great extent decentralized. This facilitates efficient and cooperative content management	OWL
Personalization	Personalized Courses are designed keeping in the students' needs and learning goals, thereby the course designed for students will differ	Student details are represented as ontology and are used by the personalization agent to search for the required information. Ontology matching is performed between learning concepts and students profile	OWL
Adaptivity	Dynamic Course content changes continuously based on the user inputs, new practices, business constraints, and experiences	Semantic web supports knowledge to be represented in different formats, which supports the varying comfort levels of the students	OWL, RDF

3 Related Work

Many researchers have proposed learning model for E-Learning system based on several parameters to serve the goal of E-Learning. Work on using Semantic Web as the next step in the evolution of E-Learning was presented by Beydoun et al. [3], highlighted the use of modular semantic-driven and service-based inter-operatibilty framework for sharing and reusing educational knowledge content for educational Semantic Web. Alsultanny [4] designed an E-Learning system. The resource format used in the Semantic Web is used for the automatic generation of hypertext using the distributed metadata in the designed system. Wu et al. [5] proposed a Semantic Web-based recommendation framework for educational resources. The knowledge in the system is represented in the form of domain ontology and user portfolio. The semantic inference is carried out for the generation of personalized recommendations based on rules representing the pedagogics.

4 SemELearn: A Semantic Web Based E-Learning System

SemELearn, a Semantic Web-based E-Learning system, is proposed in this paper, and this system has four modules: student profile, instructor, inference, and knowledge base as shown in Fig. 1.

The student profile module covers the registration process for the students with an intake of information regarding learning requirements and the understanding level of the student. Information captured by the student profile module is represented in the form of ontologies. When profile information is represented as ontologies, it contributes as an additional advantage. The first advantage is a validation of information represented by the ontologies can be performed by using reasoners, and further additional characteristics about students can be inferred by utilizing the reasoning support offered by the ontologies, and ontologies support the reuse and sharing of information.

Instructor module creates the content of courses offered in E-Learning, along with instructors' profile which contains the details of instructor-like qualification, interest areas, and experience; basically, this module generates the resources that will be consumed by the students. Instructor is the person who is going to create online digital learning resources which are commonly called as learning objects in E-Learning system. These learning objects can be in any form like text, audio, video, article, book, and chapters. Individuals are created in the subject ontology for each learning object added by the instructor.

An expert in the domain is appointed for creation of subject ontology. Domain experts create the subject ontology which contains the detail categorization of subjects and is stored along with the student profile ontology and inference rules

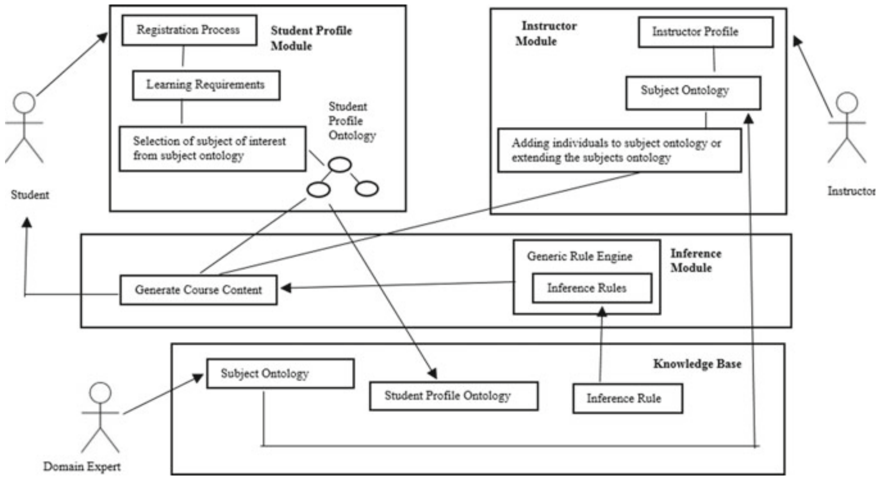


Fig. 1 SemELearn system

in the knowledge base, which is the repository for all ontologies and inference data needed to perform reasoning.

Inference module performs the reasoning on the ontologies. Reasoners are specialized software used for inferring facts from a set of explicitly asserted facts represented as ontologies. In this paper, a generic rule engine is used which works on the user specified rules. Rules contain the logic for recommending the required information regarding the course to the students.

Although many researchers have proposed the Semantic Web-based E-Learning system, they have not emphasized the role that can be played by keeping a common subject ontology for the instructors and students, the subject ontology can categorize the subject correctly without any ambiguities and is constructed by taking the service of domain experts. The other aspect which is novel in this paper is the use of reasoners based on rules. Rules represent the logic of generating personalized content for the students. The rules are written in the Jena Semantic Web development framework and are easy to interpret and understand and can be modified easily to adapt to the changing learning requirements of the students.

5 Implementation

To implement the proposed system, a Semantic Web application is created. Semantic Web application comprises of some discrete components like statements, uniform resource identifier (URI), Semantic Web languages, and ontology and instance data.

Statements are represented as tuples and consist of subject, predicate, and object. They define the information structure and are linked together to represent semantic relationship. URI gives distinctive name to every component of statement to evade naming clashes across WWW. Semantic Web language is used to express the statements. Keywords in the language are used to provide instruction to Semantic Web tools. Ontology consists of statements that describe concepts, relationships, and restrictions. Instance data comprises of statements about specific individuals of concepts, and it is built on the ontology.

Tools are required for creating Semantic Web application. Semantic Web environment contains the tools required for constructing Semantic Web application. Semantic Web environment consists of ontology editing tools, ontology reasoners, compiling and execution tools, code editing tools, and Semantic Web programming framework. For Java code compilation and execution, Java SDK is used. Code editing is carried out by Eclipse integrated development environment. For creating and editing ontologies, ontology editors are used. Protégé ontology editor will be used for implementing the proposed system. Generic rule engine and description logic reasoners perform reasoning on ontologies. For programming ontologies, Semantic Web programming framework is used, and Jena Semantic Web application development framework will be used for implementing the proposed framework. Rules will be written in Jena Semantic Web framework format.

6 Conclusions

This paper is started with the concept of E-Learning; next the significance of E-Learning and the growth of E-Learning are discussed. The advantages of replacing the World Wide Web with the Semantic Web for E-Learning are emphasized, and the most prominent advantage, in the Semantic Web machines, not only understands the information displayed but also can derive inferences to make intelligent decisions based on the available facts. The characteristic feature of the Semantic Web is highlighted next. The SemELearn system is proposed next. The architecture of the SemELearn system is depicted followed by the implementation details. In future work, the SemELearn system will be implemented using the tools discussed, and further study will be carried out on the modification which can enhance the performance of the proposed system.

References

1. Cross J (2004) An informal history of eLearning. On Horiz
2. Shah NK (2012) E-learning and semantic web. Int J e-Education, e-Business, e-Management e-Learning 2(2):113
3. Beydoun G (2009) Formal concept analysis for an e-learning semantic web. Expert Syst Appl 36(8):10952–10961
4. Alsultanny YA (2006) E-learning system overview based on semantic web. Electron J e-Learning 4(2):111–118
5. Wu L, Liu Q, Zhou W, Mao G, Huang J, Huang H (2020) A semantic web-based recommendation framework of educational resources in E-learning. Technol Knowl Learn 25(4):811–833

A Survey on Collaborative and Effective Implementation of Augmented Reality (AR) Interactive Learning Application

Satya Kiranmai Tadepalli, Fariyal Ajrad, Nimisha Malreddy, and B. Medha Reddy

Abstract As schools seek interactive learning methods to make complex concepts easily graspable, it is observed that technology can be leveraged toward the same. In particular, augmented reality has the potential to make demanding topics not seem like rocket science. With real-time visualization from day-to-day devices like smartphones and tablets, regular pedagogical strategies can be modified for the betterment of both students and teachers. This AR interactive learning application is specially designed for the NCERT science textbook of X standard CBSE board. Diagrams from the book can be scanned and visualized in 3D. The Unity game engine and Vuforia AR SDK are used for the development of the app. These models can be zoomed in and out, rotated, and rigged for an immersive experience. Along with the 3D model, a dialog box consisting of a brief description, additional resources, and aided text-to-speech audio functionality is available. The models and other assets are downloaded from TurboSquid. Additionally, to test their knowledge, students can participate in timed quizzes which have questions from the NCERT textbook for the particular topic.

Keywords Augmented reality · NCERT biology

S. K. Tadepalli (✉) · F. Ajrad · N. Malreddy · B. M. Reddy
Department of Information Technology, Chaitanya Bharathi Institute of Technology(A), Gandipet, Hyderabad, Telangana, India
e-mail: tskiranmai_it@cbit.ac.in

F. Ajrad
e-mail: ugs19127_it.fariyal@cbit.ac.in

N. Malreddy
e-mail: ugs19137_it.nimisha@cbit.ac.in

B. M. Reddy
e-mail: ugs19135_it.medha@cbit.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_63

1 Introduction

Augmented reality (AR), in simple words, can be described as integration of the users' real-world environment with computer-generated or digital content. Virtual reality (VR) on the other hand gives an opportunity for a more immersive experience by creating a full-fledged artificial environment that replaces the users' reality, unlike augmented reality [1]. In order to view digital content in a real-time environment, AR requires a device (smartphone, tablet, data glasses, etc.) with a camera and AR software. Computer vision permits AR systems to extrapolate virtual content to appropriate locations or targets. It then realistically presents the digital content in a real-world environment through display devices. AR can be categorized into two types as: Marker-based (Augmented reality applications are set off by using a set of predetermined physical images known as markers that are captured by the cameras to place digital media on them [2]) and Markerless AR (Independent of markers and gives us the freedom to decide where to display the digital elements [2]).

2 Augmented Reality in Education

It has a variety of applications across different sectors of society. However, it has shown its true potential in the field of education. Technology in education influences and motivates students to learn actively, leading to effective learning processes [3]. AR has proven to be of great potential in making the educational experiences more agile, effective, and purposeful. Not to mention, the integration of AR and education has been in the radar of extensive research as it allows students to get lost in realistic experiences [3]. Applications of AR in education are found in many fields, including biology, astronomy, chemistry, mathematics, history, geography, and physics.

3 Literature Survey

AR is currently a technology medium that offers a unique opportunity to combine the physical and virtual worlds. AR (as a learning aid) was used in a study to compare and evaluate students studying about microorganisms and found that the method actually improves biology learning [4]. In addition, the AR Marine Scientific program aims to raise awareness of their marine environment among students. The application was found to help improve the learning of their underperforming students. A recent study on virtual and augmented reality effects on K-12, higher and tertiary education students' twenty-first century skills found that the use of AR in the classroom improved levels of concept comprehension, creativity, motivation, retention, and collaboration [5]. Many AR applications also include game elements (gamification) to grab the learner's attention (Table 1).

Table 1 Literature survey table

Paper title	Author	Methodology	Limitations/Conclusion
AR learning platform for mentally differently abled children (2021)	Geetha et al. [6]	– Primary playbook with alphabets and rhymes-Unity 3D, Vuforia AR SDK, Paint 3D	– Audio-assisted contents can be helpful for differently abled children
Microorganisms: integrating augmented reality and gamification in a learning tool (2021)	Ramli et al. [4]	– Includes AR and gaming features – Unity 3D and Adobe Photoshop, AutoDesk 3D Max	– Limited to only microorganisms – Available only on Android platform
The integration of augmented reality (AR) in learning environment (2020)	Zambri and Kamaruzaman [7]	– AR-based learning app based on ARCS and CTML theories for Malaysian education system	– No use of 3D assets. Only AR multimedia such as videos are used. This is available on video streaming apps too
Creating an augmented book from a geography textbook (2019)	Vahldick and Bittencourt [8]	– App to study geography of Africa for IX class students	– Lack of information for the models used, i.e., no text/audio descriptions
Real time 3D magnetic field visualization based on augmented reality (2019)	Liu et al. [9]	– Provides real-time visualization of 3D magnetic field using AR	– When the number of magnetic lines calculated increases, the model is not rendered as expected
Innovations in tourism industry and development using AR and VR reality	Katkuri et al. [10]	– Tourist guide that gives information like history and explorations about a few tourist attractions via means of multimedia	– Does not use GPS to track the users' location, so that it can send relevant notifications automatically
Extending a user involvement tool with virtual and augmented reality. (2019)	Florea et al. [11]	A survey that studies the effect in user engagement in a living laboratory by augmented and virtual reality tools	– VR clients were experienced as innovative, easy to use, interesting, and fun, whereas AR client to be playful and empowering
Integration of augmented reality in the teaching of English as a foreign language in early childhood education	Redondo et al. [12]	This study presents the development and evaluation of an educational experiment related to early childhood education, with a special focus on the learning of EFL	– The interaction with AR in the early childhood education classroom creates a distinct atmosphere in which pupils improve their socio-affective relationships

Table 2 Tools and technologies used

Tool	Platform	Supported features
Vuforia	Android, iOS, UPW, and Unity Editor	Unity, 3D recognition, geolocation, cloud storage, and smart glass
Unity	Windows, Linux, and Mac	3D world building, AR, VR, gameplay, cinematic studio, and engineering feature set

4 Tools and Technologies

The software tools needed to build this AR application are Unity and Vuforia SDK. Unity is a game engine that uses the 3D models designed to create the application with real-time rendering [8]. Vuforia is a software development kit (SDK) that supports different operating systems [8]. 3D models can be designed using Adobe Photoshop or external digital media software like TurboSquid. This is given in Table 2.

5 Methodology

The application logic mainly resides in a database of many AR markers, each acting as a distinctive label for a set of input images. The main framework will be implemented using Unity 3D, and also SDK of Vuforia would be incorporated. 3D models are acquired from various open-source websites and Photoshop for designing AR markers. The procedure is primarily divided into three main steps.

5.1 Set up the Required Software

In this project, the first step is to set up the Unity game engine. This step involves creating a Unity project and selecting the platform to work on—Android, IOS, etc. The Android 6.0 'Marshmallow' (API Level 23) is highly preferred since most of the Android mobile phones manufactured from 2017 and above work on that level of software. This step also involves setting up the Vuforia engine.

5.2 Create User Interfaces

Next step is to create the user interfaces (UI) for the application. These UIs can be made on designing tools like Figma, AdobeXD, Adobe Illustrator, etc. The main features in the interface include a home page, a dialog box containing a description of the biology textbook diagrams, an audio button which will be used to describe the text content in the dialog box, the quiz section interface, and other basic buttons. The buttons are coded using C# script to perform their functionality.

5.3 Importing Assets

The third main step is to create marker images. Marker images, in this case, will be the diagram pictures in the biology textbook. The images are added into a database in Vuforia engine, and each image is given a rating. A rating of 4 and above out of 5 is required for the image to qualify and be used. The database is created and downloaded to be imported into the Unity game engine. Three-dimensional model assets are then imported into the asset library in Unity engine and then placed accordingly on marker images to preview how they will look when it appears on the mobile screen.

6 System Architecture

The capturing module captures the image from the camera. It takes the video from the live video feed which is divided by frames. Those frames are going to the image processing module where the marker is detected and tracked. Create a binary image, a digital image with two possible values for each pixel. The two colors commonly used in binary images are black and white. These binary images are provided as inputs to the image processing engine. Then, the tracking module calculates the pose (six degrees of freedom), i.e., the 3D location and orientation of an object in real time.

Next, it is passed to the rendering module which combines the digital object with the real-world visual to generate the augmented image. Finally, it renders the augmented digital object on the image (Fig. 1).

7 Conclusion

Making conventional pedagogy fun proves more effective in students grasping concepts. As young minds are very fickle, they are resilient toward focus. So, we made a platform where they can both interact and learn. Practical constraints such as a lack of microscope or feasibility of procuring certain elements in biology lessons hinder students' curiosity in visualizing the concepts. Traditional textbooks make it difficult to comprehend by comparing and contrasting. Our app solves this problem by providing a crisp 3D model of the object that can be panned, rotated, and viewed from any angle. Teachers and parents have a hard time teaching their children. This

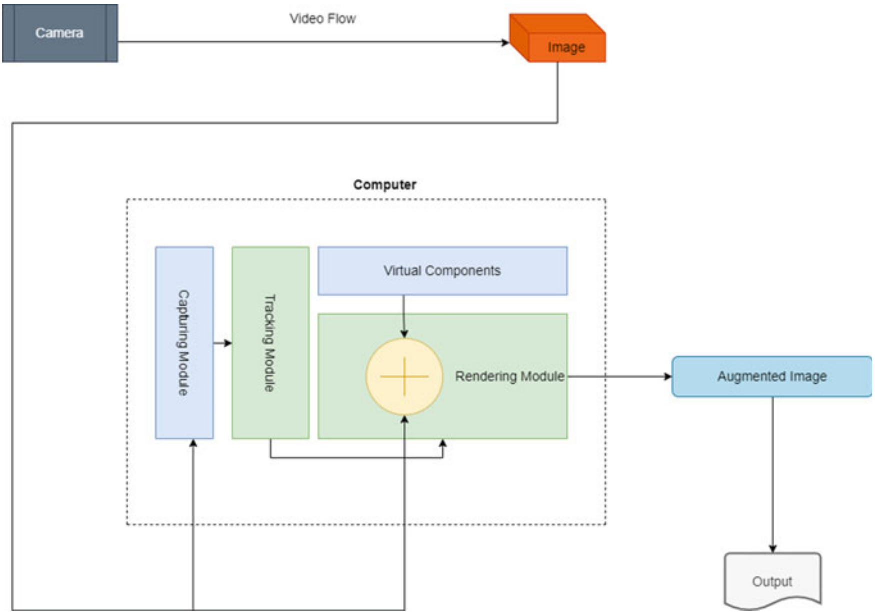


Fig. 1 System architecture

easy-to-use app for parents, teachers, and students, with simple but effective UI-UX, solves the hassle of aforementioned struggles.

References

1. https://blog.logrocket.com/create-ar-experience-unity-vuforia/
2. https://www.cgdirector.com/unity-system-requirements/
3. Saidin NF, Halim NDA, Yahaya N (2015) A review of research on augmented reality in education: advantages and applications. Int Educ Stud 8(13):1–8
4. Ramli RZ, Marobi NAU, Ashaari NS (2021) Microorganisms: integrating augmented reality and gamification in a learning tool. Int J Adv Comput Sci Appl 12(6)
5. Papanastasiou G, Drigas A, Skianis C, Lytras M, Papanastasiou E (2019) Virtual and augmented reality effects on K-12, higher and tertiary education students' twenty-first century skills. Virtual Reality 23(4):425–436. https://research.aimultiple.com/augmented-reality/
6. Geetha S, Anbarasi LJ, Raj BE, Gupta A, Vardhan Prasad A (2020) AR learning platform for mentally differently abled children. In: Seventh international conference on information technology trends (ITT). IEEE, pp 180–184
7. Zambri AA, Kamaruzaman MF (2020) The integration of augmented reality (AR) in learning environment. In: Sixth international conference on e-Learning (econf). IEEE, pp 194–198
8. Vahldick A, Bittencourt DL (2019) Creating an augmented book from a geography textbook. In: IEEE 19th international conference on advanced learning technologies (ICALT), vol 2161. IEEE, pp 347–348
9. Liu X, Liu Y, Wang Y (2019) Real time 3d magnetic field visualization based on augmented reality. In: IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 1052–1053

10. Katkuri PK, Mantri A, Anireddy S (2019) Innovations in tourism industry and development using augmented reality (AR), virtual reality (VR). In: TENCON 2019–2019 IEEE region 10 conference (TENCON). IEEE, pp 2578–2581
11. Florea C, Alavesa P, Arhippainen L, Pouke M, Huang W, Haukipuro L, Väinämö S et al (2019) Extending a user involvement tool with virtual and augmented reality. In: IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, pp 925–926
12. Redondo B, Cózar-Gutiérrez R, Sánchez Ruiz R (2020) Integration of augmented reality in the teaching of English as a foreign language in early childhood education. Early Childhood Educ J 48(2):147–155

Novel Ill-Defined Based MCDM Technique to Make Effective QoS Using Cloud Service Selection

V. N. V. L. S. Swathi, G. Senthil Kumar, and A. Vani Vathsala

Abstract In the recent era, cloud service plays a significant role for various cloud-based applications in order to make effective services and improve the quality of service (QoS). Based on the increase in user requirements, the services on the cloud are being utilized. For the utilization of services, an effective methodology is needed to rank the services to improve the QoS. Here in the cloud services, ranking the services is difficult as it depends on multi-criteria. In this case, the decision making helps in making an effective decision based on preference. In this paper, a novel ill-defined based MCDM framework is proposed to make an effective process of selecting the services being utilized in order to improve the QoS. The problem of overcoming the ambiguity in capturing and handling the decision making cloud based services. Then enhanced MCDM method is proposed to overcome the problem of rank reversal by considering the multiple QoS attributes to make improvements in ranking the cloud provider and prioritizing them. In the performance analysis, the service accessibility, data reliability, and cost are determined as the effectiveness of the proposed method is identified. The proposed system is compared with various other existing methods such as TOPSIS-based cloud selection and AHP + TOPSIS-based cloud selection.

Keywords Cloud service · Fuzzy · Quality of service · Decision making

V. N. V. L. S. Swathi (✉)
Department of Computer Science and Engineering, SRM Institute of Technology and Sciences, Chennai, India
e-mail: vs9189@srmist.edu.in; vnvls.swathi@cvr.ac.in

Department of Computer Science and Engineering, CVR College of Engineering, Hyderabad, India

G. Senthil Kumar · A. Vani Vathsala
Department of Computational Intelligence, SRM Institute of Technology and Sciences, Chennai, India
e-mail: senthilg1@srmist.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_64

1 Introduction

As generally, cloud service plays a significant in orchestrating the services while dealing with various business process modules related to cloud services. For cloud services, services can be advertised, locate the services, and invoke through remote access via the Internet. While processing the services, several XML schemas, i.e., WSDL, SOAP, etc., are applied to initiate to construct the services in a loosely coupled way for certain applications as they satisfy certain requirements based on functional and non-functional characteristics [1]. In recent days, new tools are developed for performing service monitoring, managing those services, and then optimizing them by constructing the application based on a cloud framework. As there are several services utilized, infrastructure as a service (IaaS), platform as a service (PaaS), software as a service (SaaS), and function as a service (FaaS) are utilized based on the input of data being processed [2]. As those data are generated using the Internet of Things (IoT) and then stores in the data repository, the stored data are utilized by the users for their various processes and features. According to the electronic health record (EHR), the health system is developed in order to manage those services as they should follow some health policies. Based on the user, i.e., patients, services should be customized based on their requirements [3].

Based on the increase in the population in 2022 as it increases drastically from 1980 till now, this growth gives tremendous services to the health care domain. Increase in the population makes the user to customize the system based on their needs. In the aspects of the cloud platform, task selection based on priority is the major thing for service-oriented features. The service-oriented features as they are agile the data and maintain those things for various service applications like electronic healthcare, supply chain management, etc., that service-oriented architecture platform makes certain analyses with respect to certain attributes. The system platform ability can be deployed, and it is evaluated based on QoS. Based on this platform, the service-oriented architecture deployment, design, and implementation can be done [4]. Then, the SOA performance can be measured with certain results. Service-oriented architecture (SOA) plays a significant role in providing services for various applications. This service platform makes the certain process of appealing the services based on the user requirement and deploying them. The services contain various processes, sub-services, and rules to be associated with certain applications. Mainly, SOA is highlighted in various IT organizations as various services are associated. There are three significant elements which are attached to the process of service in SOA [5].

1. For certain applications, reusability has to be applied.
2. Ability to communicate among the modules or other external applications.
3. Allow the application to apply the future needs based on a certain organization.

The main problem of scheduling and making some tasks priority is cloud computing which helps to define the architecture, which is service oriented. Then, architecture should be defined based on the user needs and organization scenario, and the organization may be electronic health care, managing supply chain [6]. Then, the SOA architecture may be analyzed more consistently with certain attributes to analyze the performance related to application specific. The application architecture makes some advantages and disadvantages to analyze the level of QoS. To the existing architecture, some architecture makes modifications with few attributes in order to optimize and configure the architecture based on the applications. Service-oriented architecture helps to exchange the communication among the modules in the system specific to applications [7]. Here, the data association makes some process of synchronization based on the features of heterogeneity. Those systems are designed based on the characteristics of service and classified into conceptual view: The process workflow is demonstrated based on the association of various process activities, logical view: Based on the services offered by the system, all the logical aspects are selected based on the previous factor, and physical view: Based on the system, various services are offered such as presentation, application, and domain aspects based on data.

Proposed a novel framework for the cloud platform by compositing the services and formulating the problem of heterogeneity and scheduling problem. Then, the service-oriented algorithm is proposed by analyzing the properties of attributes in order to ensure QoS. Analysis based on the composition of services makes the algorithm to better perform than the existing methods, SCA and ADAPT. In this paper, a novel ill-defined based MCDM framework is proposed to make an effective process of selecting the services being utilized in order to improve the QoS. The problem of overcoming the ambiguity in capturing and handling the decision making cloud based services. Then, enhanced MCDM method is proposed to overcome the problem of rank reversal by considering the multiple QoS attributes to make improvements in ranking the cloud provider and prioritizing them.

The paper is organized as follows: existing literature articles are discussed in Sect. 2, then problem definition in Sect. 3, and proposed methodology in Sect. 4. Then, performance analysis is done on Sect. 5 and conclusion in Sect. 6.

2 Literature Survey

Over the past few years, the cloud service provider is increasing drastically as it considers the problem of cloud service selection in order to improve the QoS. Lu et al. [8] has proposed a framework as it can access the cloud services of three IaaS as it is denoted as SMICLOUD, and it is compared with other existing frameworks. In this framework, AHP method is deployed in order to determine the criteria weight of the cloud services. The three IaaS cloud services are analyzed, and the cloud services are accessed. The three IaaS service providers are mentioned as problem decomposition, evaluating the cloud selection based on the priority and rank IaaS service providers.

1. For the link selection, the hierarchical structure is constructed based on the QoS attributes and service provider.

2. To calculate the criteria weight, utilize the pairwise comparison matrix.
3. Using the criteria weight, IaaS cloud service providers are ranked.

In the analysis, there are several key performance indicators (KPIs), and there are various cloud service providers that compare them. Then, enable the criteria weight measurement based on user preference and metrics interdependencies. Masdari and Khezri [9] uses AHP-based methodology that helps to identify the cloud service provider database, and it makes three criteria and seven sub-criteria. Cloud Genius framework is proposed with the integration of genetic algorithm as it asses the IaaS cloud services. In this framework, the best IaaS service provider is selected based on various QoS parameters. There are several works which are deployed to overcome the problem of uncertainty in order to perform cloud selection, and it is associated with fuzzy concept [10–12]. The fuzzy-based AHP framework is proposed during the quality assessment in order to handle the quality of vagueness and subjectivity. The performance factor improves the quality of experience (QoE) based on feature model. Then the, fuzzy-based AHP method [13] helps to analyze the various cloud services with multiple QoS attributes.

To overcome the problem of cloud service selection, the associated framework of fuzzy logic and AHP is proposed [14–17]. The cloud service is reusable as it gets contributed based on two aspects such as service provider and customer. Then, for the above scenario, MCDM-based cloud service integrated with fuzzy is proposed as it used SAW concept into it. Then, [18] fuzzy SAW model with VIKOR model is deployed as it proposed the trust-based fuzzy MCDM hybrid method, and in order to ensure the trust ability, there are 15 factors which are added to measure the degree of the cloud service.

3 Problem Definition

As there are multiple cloud service providers available in the recent era, customers can deploy any decision-making methods to analyze their decision by analyzing and evaluating the provider rank based on the user requirement. There are several existing decision-making methods available to analyze and rank the cloud provider as they are specific to multi-attribute decision making (MADM) method. The main problem is the better cloud provider as it can be deployed based on the decision-making approach. To make the provider better, ranking and decision problem is considered as it is specific to multi-criteria attributes. Better multi-criteria decision making is formulated by addressing the problem of time-aware trust-based service selection and ranking the service list to be considered.

4 Proposed Methodology

In this novel ill-defined based MCDM framework as represented in Fig. 1, there are three elements are contributed as listed below.

(a) Constructing the cloud broker with the association of discovering the cloud service and ranking the cloud service
(b) The standard for the cloud service provider based on monitoring and auditing the service

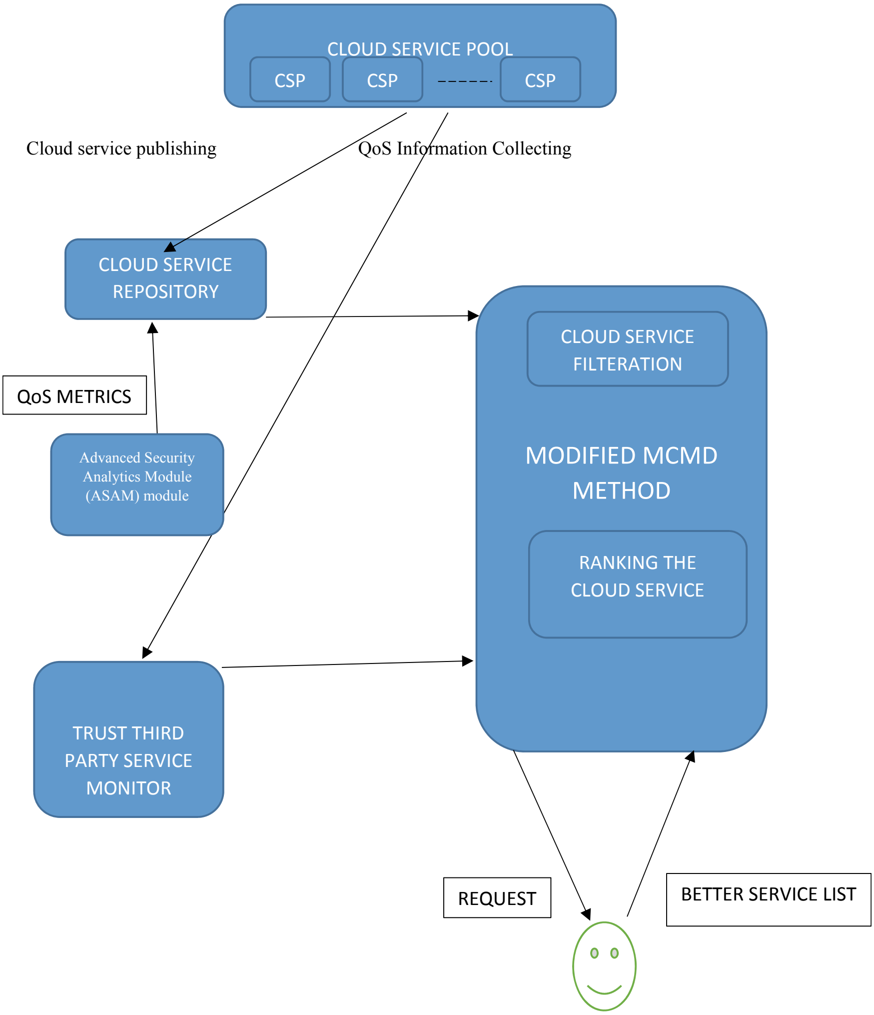


Fig. 1 Ill-defined based MCDM framework

(c) Repository to store the information based on the cloud service.

A **Constructing the cloud broker with the association of discovering the cloud service and ranking the cloud service**
Framework deployed based on the cloud broker as represented in Fig. 2 included multiple activities to obtain cloud service discovery, and rank the cloud service. To perform the process of ranking and discovery, it has interacted with the cloud repository as it filters the improved cloud provider as it is based on the user requirement. Based on the QoS performance parameters by the cloud users, the cloud broker will perform a ranking to filter the improved cloud service provider. The ranking is performed for each cloud service established by the provider. The main use of the cloud service discovery module is to discover and rank the cloud information based on the services provided by the cloud.

B **The standard for the cloud service provider based on monitoring and auditing the service**
This cloud service standardization audits and monitors the activity of cloud service periodically in a regular time interval. The cloud service tests are done based on QoS criteria such as cloud availability, reliability, throughput, and efficiency. Then, the testing process makes QoS criteria based on the cloud provider, and then the cloud standard is made based on the cloud services, and its repository is analyzed based on a regular basis and then stored in the repository of the cloud service.

C **Repository to store the information based on the cloud service.**
The cloud repository acts as the database, which stores the information, and it contains the cloud service provider information and its QoS attributes. The data is stored in the repository, which contains cloud provider information and monitoring of the cloud services based on the third party, and those performances are stored in the repository. The database is used for scanning the services based on a cloud broker, and it makes candidate services based on the customer needs. The trust-based cloud information is ensured based on the advanced security analytics module (ASAM) module.

(i) **Advanced security analytics module (ASAM) module**

This advanced module makes the flow network ensure security to perform data analytics based on anomaly detection tool, which helps to detect the intrusion of the zero-day network as it uses continuous stream mining engine technology. It classifies network intrusion as challenging the network security in the real-time application. It offers better intelligence in order to identify the wide spectrum to ensure internal and external security as it is represented in Fig. 2.

The ASAM security makes the problem to identify the bad link on source to destination, distributed denial of service (DDoS), and identify the suspected flow.

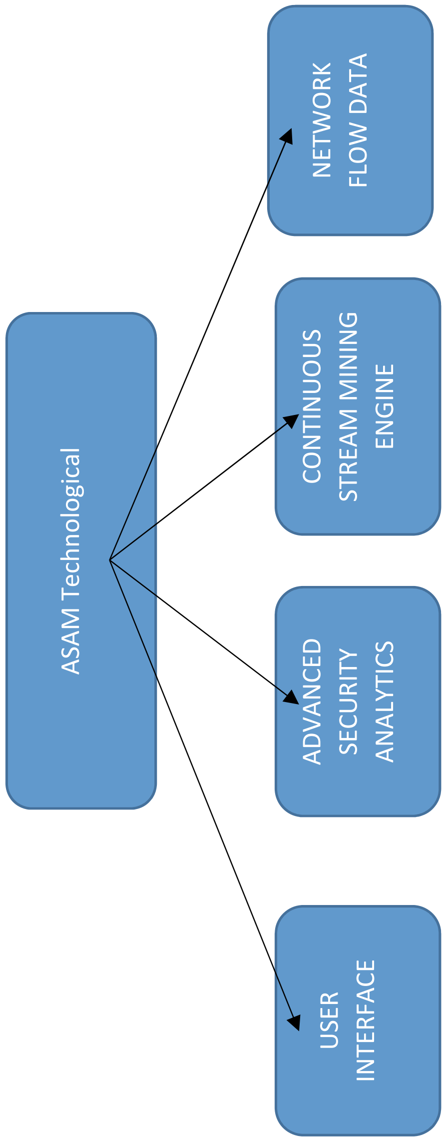


Fig. 2 ASAM significance

This ASAM offers better value data robust, scalability, and data proven by monitoring the network bandwidth and analysis the network traffic.

(ii) **Measure the cloud service index**

For analysis of the cloud provider, there is a service index to be measured as it helps to measure the services provided by the cloud. In the initial phase of the framework as denoted in Fig. 2, it makes several cluster groups and each group considers multiple criteria as mentioned below,

(a) Accountability metric helps to measure the properties related to the services provided by the cloud application.
(b) Cloud agility helps to determine the cloud service impact as it makes the user ability by applying the strategy to reduce network disruption.
(c) Data assurance indicates the degree of the cloud service based on the service availability.
(d) Criteria based on finance help to identify the amount spent on the particular service based on the user requirement.
(e) Security and privacy make to identify the service provider with improved effectiveness in order to control the service accessibility and their service data.

The above criteria make the service level agreement (SLA)-based on certain QoS rules and standards.

(iii) Modified MCDM (TOPSIS) Algorithm

As there are various decision-making algorithms that exist, it creates various subgroups in terms of cost and benefits. The technique for order performance by similarity to ideal solution (TOPSIS) is proposed for the rationally logic and easy computational process. Then, it makes various possible criteria for finding the best alternative solution, and weight metrics are also added. In the existing TOPSIS, the problem of rank reversal occurs as it is represented in various MCDM methods such as AHP, TOPSIS, ELECTRE, or PROMETHEE. In the alternative process of finding the best criteria by adding or removing operations from the set of candidates available in the information. In this method, rank reversal is done by the Euclidean distance.

In the modified TOPSIS, an effective Minkowski distance metric is formulated to perform the rank reversal process by varying the P value.

P occurs between two points $p = \{p_1, p_2, \ldots p_n\}$ which belongs to rational R^n and $q = \{q_1, q_2, \ldots q_n\}$ which belongs to rational R^n.

$$p(x, y) = \left(\sum_{i=1}^{n} p_i - q_i^r \right)^{1/r}$$

In this distance formulation, $p = 1$ is used for this method TOPSIS.

Algorithm 1: Enhanced MCMN Algorithm

Input: Set of Criteria CR.

Output:

1. Perform identification and selection of criteria and sub-criteria.
   Set of Criteria CR = {$\mathbf{CR_1}$, $\mathbf{CR_2}$, … $\mathbf{CR_n}$}
   Based on the criteria, cost and its benefits are maximized/minimized
2. To perform alternative criteria AC,
   Set of alternative criteria AC = {$\mathbf{AC_1}$, $\mathbf{AC_2}$, … $\mathbf{AC_m}$}
3. Based on the CR and AC, formulating the evaluation matrix by constructing $m \times n$ matrix as,

$$E = (e_{ij})$$

   where $i = \{1, 2, 3, \dots m\}; j = \{1, 2, 3, \dots n\}$
4. Based on input data fetched from Step 1 to 3,
   Criteria weight (CW) = $(\mathrm{CW}_j), j = 1, 2, 3, \dots n.$
   Based on the cost 'T' criteria, criteria can be maximized/minimized
5. Normalize E-matrix.

$$\overline{E} = (\overline{e_{ij}})$$

   where $i = \{1, 2, 3, \dots m\}; j = \{1, 2, 3, \dots n\}$
   Then, normalize the entries of the matrix.
6. Based on CW criteria weight, prioritized the criteria.
   Then, the weight of each criteria is calculated based on

   (a) Mean weight

$$a_{i,j} = \frac{1}{a_{ji}}, a_{ij} > 0$$

   (b) Weight entropy
   The entropy-based objective weighting technique is the one that was utilized. A fairly well-liked method of objective weighting is entropy. Traditional entropy-based weighting strategies, on the other hand, presuppose that each QoS component has a set value, whereas in our approach, each QoS aspect is allowed to employ unknown QoS values (QoS intervals). For every QoS factor and its relative weight, we calculate the entropy.

$$\text{Alternative}(A) = \{A_1, A_2, A_3, \dots A_n\}$$

$$\text{Decision Criteria } C = \{C_1, C_2, C_3, \dots C_n\}$$

$$\text{Entropy } e_j = -\frac{1}{\ln m}\sum_{i=1}^{m} P_{ij}.\ln P_{ij}$$

$$P_{ij} = \frac{g_{ij}}{\sum_{i=1}^{m} g_{ij}}$$

$$\text{where } 0 \ln 0 = 0;$$
$$i = 1, 2, \ldots m$$
$$j = 1, 2, \ldots n$$
$$g_{ij} \rightarrow \text{decision matrix } m \times n$$

(c) Variation weight based on co-efficient

$$V(W) = \beta W(0)[W] + (1-\beta)W(S)[W] \text{ where } 0 \leq \beta \leq 1$$

6. Calculate the positive and negative ideals based on maximizing and minimizing the normalized E-matrix.
7. Alternative ranking analysis to obtain stability.

5 Performance Analysis

In the analysis, criteria and sub-criteria are mentioned, and its QoS performance metrics are represented and values are determined. Here, there are three criteria which are mentioned as listed below,

- Performance:
  Sub-criteria—Functionality and response time
- Cost: Storage, memory, and acquisition cost.
- Security and privacy:
  Access control and data integrity.

Then, the above values are analyzed based on weight variation based on co-efficient, weight entropy, and mean weight as represented in Tables 1, 2, and 3.

The number of comparisons is determined based on the variation in the criteria from 2 to10, and the proposed model performs better for each criteria as the effective value gets calculated, and it performs effectively as compared to other existing models such as AHP, fuzzy-based AHP, and improved TOPSIS. It is represented in Fig. 3.

In Fig. 4, the cloud service rank is determined based on the variation in the cloud service provider from 1 to 10. Then, the ranking value gets increased for the proposed model and its large fluctuation as the average value gets improved as compared to the existing models like AHP, fuzzy-based AHP, and improved TOPSIS.

Table 1 Weight variation based on co-efficient for enhanced MCDM method

S. No.	Criteria/Sub-criteria	Group	Benefit (max) or cost (min) criteria	Type
1	Performance	CR1	0.020	
	Functionality	CR11	Max 0.007	Qualitative
	Response time		Min 0.013	Quantitative
2	Cost	CR2	0.0691	
	Storage cost	CR21	Min 0.164	Quantitative
	Memory cost	CR22	Min 0.113	Quantitative
	Acquisition cost	CR23	Min 0.231	Quantitative
3	Security and privacy	CR3	0.091	
	Access control	CR31	Max 0.232	Qualitative
	Data integrity	CR32	Max 0.042	Qualitative

Table 2 Weight entropy for enhanced MCDM method

S. No.	Criteria/Sub-criteria	Group	Benefit (max) or cost (min) criteria	Type
1	Performance criteria	CR1	0.051	
	Functionality sub-criteria	CR11	Max 0.071	Qualitative
	Response time		Min 0.032	Quantitative
2	Cost	CR2	0.565	
	Storage cost	CR21	Min 0.187	Quantitative
	Memory cost	CR22	Min 0.197	Quantitative
	Acquisition cost	CR23	Min 0.198	Quantitative
3	Security and privacy	CR3	0.198	
	Access control	CR31	Max 0.232	Qualitative
	Data integrity	CR32	Max 0.210	Qualitative

The consistency ratio gets calculated based on the cloud service to be a selection from 0 to 9 as shown in Fig. 5. Then, the consistency gets improved based on the data for the proposed model as compared to the existing models like AHP, fuzzy-based AHP, and improved TOPSIS.

Table 3 Mean weight for enhanced MCDM method

S. No.	Criteria/Sub-criteria	Group	Benefit (max) or cost (min) criteria	Type
1	Performance criteria	CR1	0.196	
	Functionality sub-criteria	CR11	Max 0.132	Qualitative
	Response time		Min 0.132	Quantitative
2	Cost	CR2	0.297	
	Storage cost	CR21	Min 0.121	Quantitative
	Memory cost	CR22	Min 0.123	Quantitative
	Acquisition cost	CR23	Min 0.128	Quantitative
3	Security and privacy	CR3	0.196	
	Access control	CR31	Max 0.128	Qualitative
	Data integrity	CR32	Max 0.128	Qualitative

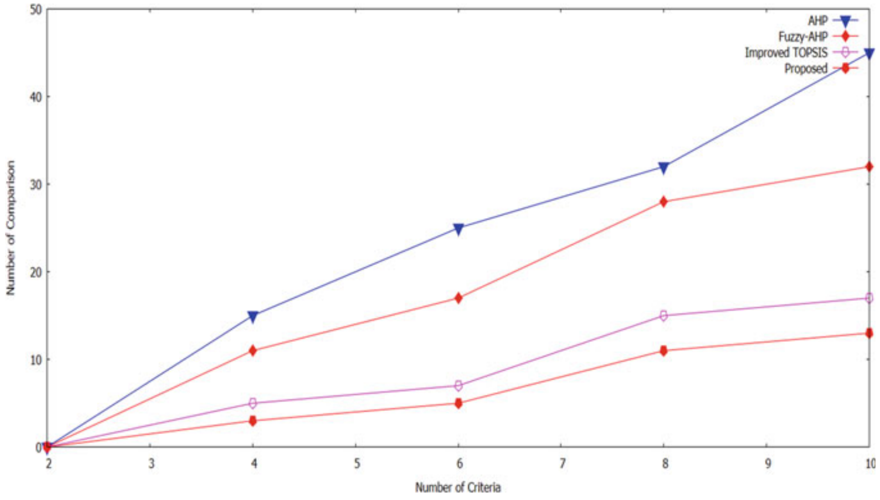


Fig. 3 Number of criteria versus number of comparisons

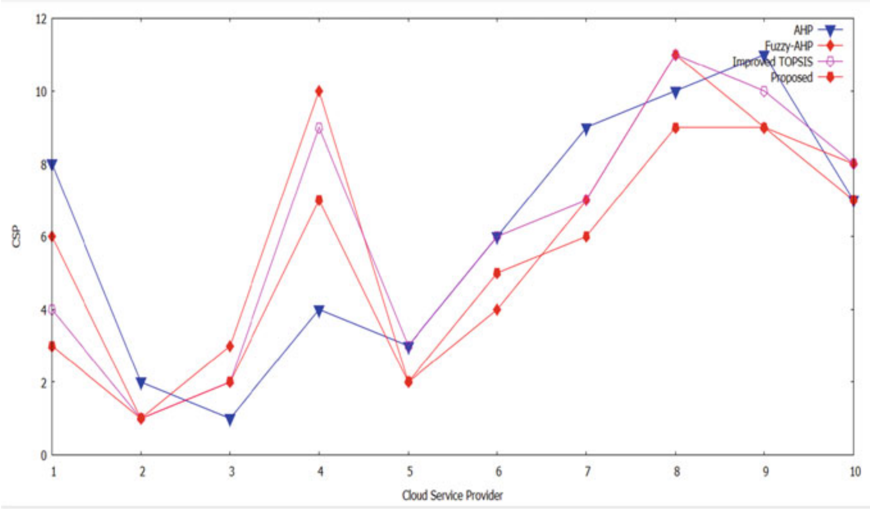


Fig. 4 Cloud service provider versus CSP (Cloud service provider rank)

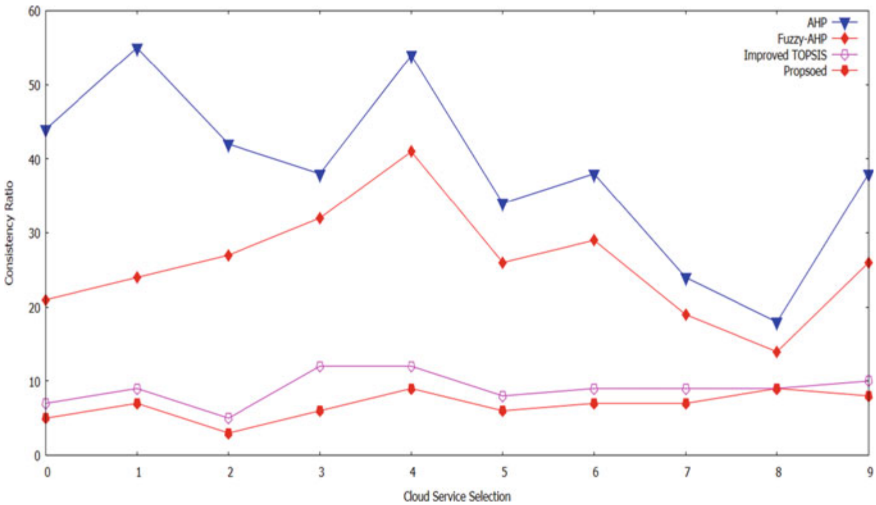


Fig. 5 Cloud service selection versus consistency ratio

6 Conclusion

Based on the problem of rank reversal, an ill-defined based MCDM framework is proposed as takes variation in P which uses Minkowski distance metric with regular interval of time. Proposed enhanced MCDM method makes an effective process of selecting the services being utilized in order to improve the QoS. The problem of overcoming the ambiguity in capturing and handling the decision making cloud based services. Then, enhanced MCDM method is proposed to overcome the problem of rank reversal by considering the multiple QoS attributes to make improvements in ranking the cloud provider and prioritizing them. Then, it takes multi-criteria-based QoS attributes such as performance criteria, cost and security, and privacy. Then, there are three weight methods such as weight variation based on co-efficient, weight entropy, and mean weight which are calculated.

References

1. Cheng DY, Chao KM, Lo CC, Tsai CF (2011) A user-centric service-oriented modeling approach. World Wide Web 14:431–459
2. Deng H, Yeh CH, Willis RJ (2000) Intercompany comparison using modified TOPSIS with objective weights. Comput Oper Res 27:963–973
3. Hwang CL, Yoon KS (1981) Multiple attribute decision making: methods and applications. Springer-Verlag Press
4. İç YT (2012) An experimental design approach using TOPSIS method for the selection of computer-integrated manufacturing technologies. Robot Comput Integr Manuf 28(2):245–256
5. Liu S, Chan FTS, Ran W (2016) Decision making for the selection of cloud vendor: an improved approach under group decision-making with integrated weights and objective/subjective attributes. Expert Syst Appl 55:37–47
6. Ma H, Hu Z, Li K, Zhang H (2016) Toward trustworthy cloud service selection: a time-aware approach using interval neutrosophic set. J Parallel Distrib Comput 96:75–94
7. Rădulescu CZ (2017) A cloud providers' services evaluation using triangular fuzzy numbers. In: Proceedings of the 21th international conference on control systems and computer science (CSCS21). pp 123–128
8. Lu Y, Zheng X, Ling L, Xu LD (2020) Pricing the cloud: A QoS-based auction Approach. Enterp Inf. Syst. 14(3):334–351. https://doi.org/10.1080/17517575.2019.1669827
9. Masdari M, Khezri H (2020) Service selection using fuzzy multi-criteria decision making: a comprehensive review. J Ambient Intell Human Comput 1–32
10. Panjun S (2020) Research on cloud computing service based on trust access control. Int J Eng Bus Manage 12:1847979019897444. https://doi.org/10.1177/1847979019897444
11. Abdel-Basset M, Mohamed M, Chang V (2018) NMCDA: a framework for evaluating cloud computing services. Futur Gener Comput Syst 86:12–29. https://doi.org/10.1016/j.future.2018.03.014
12. Alabool H, Kamil A, Arshad N, Alarabiat D (2018) Cloud service evaluation method-based multi-criteria decision-making: a systematic literature review. J Syst Softw 139:161–188. https://doi.org/10.1016/j.jss.2018.01.038
13. Liu Z, Wang X, Wang W et al (2022) An integrated TOPSIS–ORESTE-based decision-making framework for new energy investment assessment with cloud model. Comp Appl Math 41:42. https://doi.org/10.1007/s40314-021-01751-9
14. Moori A, Barekatain B, Akbari M (2022) LATOC: an enhanced load balancing algorithm based on hybrid AHP-TOPSIS and OPSO algorithms in cloud computing. J Supercomput 78:4882–4910. https://doi.org/10.1007/s11227-021-04042-6
15. Xu X, Xie J, Wang H et al (2022) Online education satisfaction assessment based on cloud model and fuzzy TOPSIS. Appl Intell. https://doi.org/10.1007/s10489-022-03289-7
16. Yang WG, Wu YJ (2019) A novel TOPSIS method based on improved grey relational analysis for multiattribute decision-making problem. Math Probl Eng 2019:1–10. https://doi.org/10.1155/2019/8761681

17. Wu S, You X, Liu H, Wang L (2020) Improving quality function deployment analysis with the cloud MULTIMOORA method. Int Trans Oper Res 27(3):1600–1621
18. Du YF, Liu D (2020) A novel approach to relative importance ratings of customer requirements in QFD based on probabilistic linguistic preferences. Fuzzy Optim Decis Ma 1–31.https://doi.org/10.1007/s10700-020-09347-4

Cluster Query Optimization Technique Using Blockchain

Dileep Kumar Kadali, D. Venkata Naga Raju, and P. Venkata Rama Raju

Abstract Blockchain is an emerging technology, and it applies to time-stamped data that can be achieved by way of a cluster of processors now not possessed with the aid of any solitary entity. Here, cluster data records are to be considered as a set of blocks. Those blocks are protected and processed respectively for the usage of cryptographic standards, namely chain. The query processing is continued until achieved exact and optimal execution is on the blocks. The query is executed after connecting queries to databases, collecting statistics from exceptional places, and showing outcomes to the machine learning application. We propose a query optimization, which is the portion of the query procedure in the database system that compares distinctive query processes and selects the only one with the least expected value in every block. A query optimizer makes one or more query tactics for every query, each of which can be a device used to run a query. The most efficient query plan is chosen and used to run the query. The experimental result has calculated the latency for each query block node.

Keywords Blockchain · Cluster · Cryptographic · Query optimization · Query optimizer · Machine learning

D. K. Kadali (✉) · D. Venkata Naga Raju · P. Venkata Rama Raju
Shri Vishnu Engineering College for Women, Bhimavaram 534202, India
e-mail: kdileepkumarit@svecw.edu.in; dileepkumarkadali@gmail.com

D. Venkata Naga Raju
e-mail: dvnraju@svecw.edu.in

P. Venkata Rama Raju
e-mail: pvramaraju@svecw.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_65

1 Introduction

Machine learning systems are used large-quantity statistics with algorithms to search for elegances predictions [1, 2]. Query optimization is similar to a completely centered method of calculating predictions for the way the device can and has to run selected queries [3]. With queries, it must not be complex methods to execute them. These combinations require large portions of computing power and complex procedures to calculate the incredible effects [4]. Query optimization studies the information in your small table and learns from the facts to make predictions approximately similar or exact to your queries [5–8]. It then makes use of the predictions to come to be aware of the most efficient methods to run operations. Even at the same time as query optimization gets a prediction incorrect, it is going to paint through the years via the usage of accumulating more information to provide more contexts. This method moreover trails the fundamental varieties of machine learning. Database management optimization refers to the circulation of techniques for taking down the database machine response time [5, 9]. Deliver directly to learn greater approximately database optimization and the specialists who manipulate databases in a prevalent variety of software industries. Databases can work fine quantities of data and petabytes of statistics [10–12]. Precise bits of records are accessed by using queries written in a specific interface language, along with SQL [13–15]. Database optimization includes making the most of the rate and performance with which records are retrieved. Database engineers, administrators, and experts work collectively to optimize machine overall concert through various methods [16, 17]. A thoughtful layout that cautiously addresses functional wishes is the inspiration for concert enhancement. The creation of queries can produce quicker effects, as well [18, 19]. Administrators and analysts are trying to find an approach to enhance their servers' records and get admission to methods and retrieval epochs through design methods, statistical evaluation, and identification of machine traffic [20–22]. Blockchain technology undertakes to make the most of the server's potential with the aid of owning a robust understanding of the structure of the information, the applications hooked up on the server, and the effect numerous duties have on the database's typical performance [23–25].

2 Query Optimization

Query optimization is the machine in which a database or records warehouse takes the entry from a user's inquiry, and revises that query to supply a brief response with few compute resources as feasible. The variety of query plan changes gains far past what a humanoid can calculate. A set of changes for the best plan makes requests quicker because it gets facts extra speedy. As soon as a query is placed, it is scanned, parsed, and validated. It is an in-house depiction of the query generated such as a query tree or a query graph. Before alternative steps of the execution, plans are made

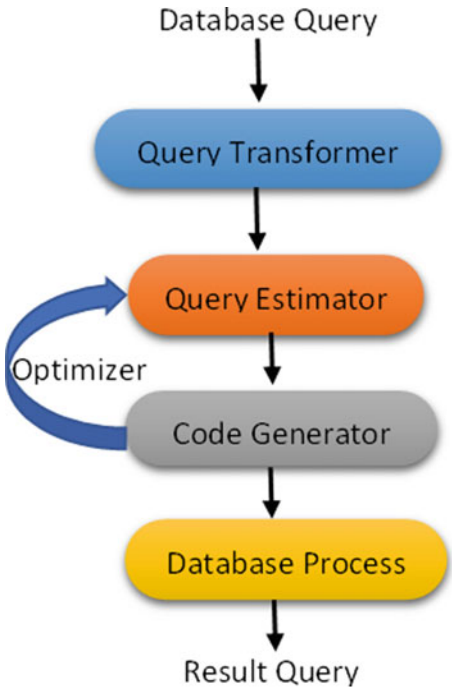


Fig. 1 Process of query optimization

for retrieving results from the database tables. This process of choosing the most suitable execution plan for query processing namely is query optimization (Fig. 1).

3 Query Optimization Using Blockchain

We can take a query for processing the query operation. Through blockchain, not only the query optimizer saves scanned data from the scanner, but also parses the data from the parser and also validation processing of the query result; so that the situation can change the perfect query process of the blockchain. The user requests in the transaction user and the database. The query is a block that will be further to a query blockchain. A query process transaction on the blockchain is an exclusive, autonomously provable, and unfalsifiable record (for example Bitcoin), therefore every user query is. By the way, the result of query blockchain is to the best ever of altogether transactions for, say, a certain query process path, or even the entire query tree that comprises every query process ever engaged.

4 Services of Query Optimization Using Blockchain

The services of query optimization allow any services to be progressive or substituted independently. This service of the user query is processed on and uses a graphical user interface running on the optimization services. The query analysis service is on query optimization and contains query processor service to use the process of querying the data storage logic. The following three services are used for query optimization.

1. The user service occupies the top level and shows information related to services available on a user interface. This service tier interconnects with other service tiers by sending results to the interface and other service tiers in the network.
2. Optimization of query logic received from the user service. It controls query operation functionality by performing detailed processing.
3. Infrastructure service is the query storage from existing databases like cloud or Big Data and others where the information is stored and retrieved. The query processing data in this service is reserved in the self-governing of query logic (Fig. 2).

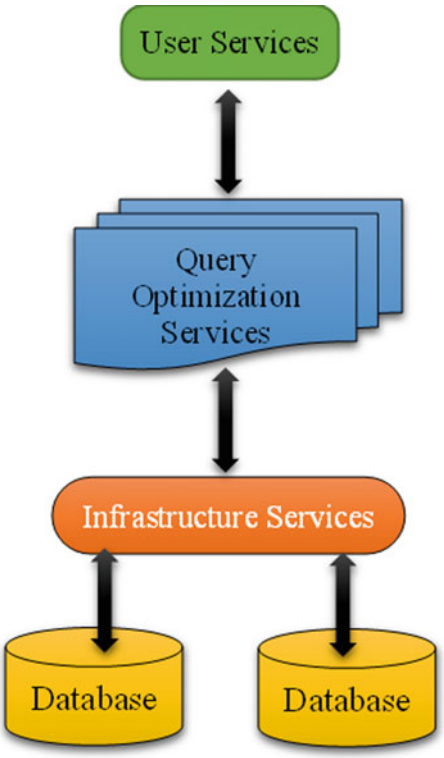


Fig. 2 Query optimization process services using blockchain

The statistical evidence for cost estimation is measured as b_r which is the number of blocks comprising tuples of r. If tuples of r are stored composed tangibly in a file, then

$$b_r = \frac{n_r}{f_r} \tag{1}$$

where

n_r is the number of tuples in a relation r.
f_r is the blocking factor of r which is the number of tuples of r that fits into one block.

5 Experimental Results

The experimental result is to measure the performance of a query optimization by means of blockchain which is hard to evaluate through esteem to only one constraint, since there are many constraints that control the performance of the blockchain environment. The evaluation starts by means of trust query nodes, as input to our arrangement to specify the number of query nodes to be reserved in each place and the duration of the experiments. Deploy and configure an isolated blockchain and set up query optimization. The column latency time is occupied through the blockchain to handle the query needs. As given in table, minimizing the number of query nodes does not have an important effect on the latency (Fig. 3; Table 1).

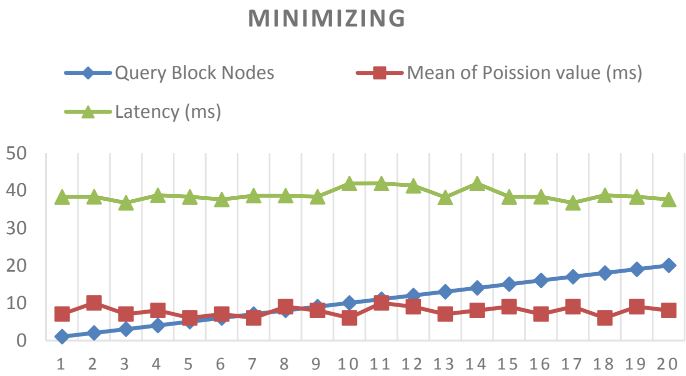


Fig. 3 Minimizing the number of query nodes does not have an important effect on the latency

Table 1 Latency and mean poisson for query block nodes

Query block nodes	Mean of Poisson value (ms)	Latency (ms)
1	7	38.319
2	10	38.348
3	7	36.72
4	8	38.73
5	6	38.319
6	7	37.61
7	6	38.634
8	9	38.634
9	8	38.348
10	6	41.872
11	10	41.9
12	9	41.3
13	7	38.18
14	8	41.872
15	9	38.319
16	7	38.348
17	9	36.72
18	6	38.73
19	9	38.319
20	8	37.61

6 Conclusion

The query optimizer, which contains this function, is a key part of the relational database and determines the most effective way to execute a given query by considering the viable query plans. Query optimization is part of the query manner wherein the database system compares specific query techniques and chooses the only one with the least predicted cost. The experimental results show that the proposed system can improve performance. After this, the use of certain query optimization techniques is improving more query processing, and query response time, and reduces the limitations of a query process system.

References

1. Singh N, Vardhan M (2018) Blockchain based E-stamp procurement system with efficient consensus mechanism and fast parallel search. J Mech Continua Math Sci 13(4):73–89. https://doi.org/10.26782/jmcms.2018.10.00007

2. Yadav AS, Shikha S, Gupta S, Kushwaha DS (2021) The efficient consensus algorithm for land record management system. IOP Conf Series Mater Sci Eng 1022(1):012090. https://doi.org/10.1088/1757-899X/1022/1/012090
3. Kadali DK, Jagan Mohan RNV, Vamsidhar Y (2012) Similarity based query optimization on map reduce using Euler angle oriented approach. Int J Sci Eng Res 3(8). ISSN: 2229-5518
4. Singh N, Vardhan M (2021) Multi-objective optimization of block size based on CPU power and network bandwidth for blockchain applications. In: Proceedings of the fourth international conference on microelectronics, computing and communication systems, Singapore, pp 69–78. https://doi.org/10.1007/978-981-15-5546-6_6
5. Kadali DK, Chandra Naik M, Jagan Mohan RNV (2022) Estimation of data parameters using cluster optimization. Springer Nature
6. Kadali DK, Jagan Mohan RNV (2014) Optimizing the duplication of cluster data for similarity process. ANU J Phys Sci 2. ISSN: 0976-0954
7. Kadali DK, Jagan Mohan RNV, Srinivasa Rao M (2016) Cluster optimization for similarity process using de-duplication. IJSRD-Int J Sci Res Dev 4(06):2321–0613. ISSN (online)
8. Jagan Mohan RNV, Kadali DK et al (2022) Crime data optimization using neutrosophic logic based game theory. Concurrency Computat Pract Exp
9. Singh N, Vardhan M (2020) Computing optimal block size for blockchain-based applications with contradictory objectives. Proc Comput Sci 171:1389–1398. https://doi.org/10.1016/j.procs.2020.04.149
10. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: Architecture, consensus, and future trends. In: 2017 IEEE international congress on big data (BigData Congress), Honolulu, HI, USA, pp 557–564. https://doi.org/10.1109/BigDataCongress.2017.85
11. Kadali DK, Jagan Mohan RNV, Chandra Naik M (2020) Unsupervised based crimes cluster data using decision tree classification. J Solid State Technol 63(5):5387–5394
12. Tschorsch F, Scheuermann B (2016) Bitcoin and beyond: a technical survey on decentralized digital currencies. IEEE Commun Surv Tutorials 18(3):2084–2123. https://doi.org/10.1109/COMST.2016.2535718
13. Singh K, Singh N, Kushwaha DS (2018) An interoperable and secure e-wallet architecture based on digital ledger technology using blockchain. In: 2018 international conference on computing, power and communication technologies (GUCON), Greater Noida, India, pp 165–169. https://doi.org/10.1109/GUCON.2018.8674919
14. Kumar A, Yadav AS, Kushwaha DS (2020) VChain: efficient blockchain based vehicular communication protocol. In: 2020 10th international conference on cloud computing, data science & engineering (confluence), Noida, India, pp 762–768. https://doi.org/10.1109/Confluence47617.2020.905780
15. Singh N, Vardhan M (2019) Digital ledger technology-based real estate transaction mechanism and its block size assessment. Int J Blockchains Cryptocurr 1(1):67–84. https://doi.org/10.1504/IJBC.2019.101853
16. Nakamoto S (2019) Bitcoin: a peer-to-peer electronic cash system. Manubot
17. Kadali DK, Jagan Mohan RNV (2022) Shortest route analysis for high level slotting using peer-to-peer. The role of IoT and blockchain: techniques and applications. CRC press
18. Singh N, Vardhan M (2019) Distributed ledger technology-based property transaction system with support for IoT devices. Int J Cloud Appl Comput (IJCAC) 9(2):60–78. https://doi.org/10.4018/IJCAC.2019040104
19. Thakur V, Doja MN, Dwivedi YK, Ahmad T, Khadanga G (2020) Land records on blockchain for implementation of land titling in India. Int J Inf Manag 52:101940. https://doi.org/10.1016/j.ijinfomgt.2019.04.013
20. Mallick S, Kushwaha DS (2012) An efficient publish-subscribe protocol for collaborative content delivery-ex-PUSH. In: 2012 international symposium on cloud and services computing, Mangalore, India, pp 152–156. https://doi.org/10.1109/ISCOS.2012.27
21. Singh N, Kumar T, Vardhan M (2019) Blockchain-based e-cheque clearing framework. Scalable Comput Pract Exp 20(3):511–526. https://doi.org/10.12694/scpe.v20i3.1506

22. Singh N, Kumar T, Vardhan M (2020) Blockchain-based e-cheque clearing framework with trust-based consensus mechanism. Cluster Comput 1–15.https://doi.org/10.1007/s10586-020-03163-6
23. Nguyen GT, Kim K (2018) A survey about consensus algorithms used in blockchain. J Inf Process Syst 14(1):101–128. https://doi.org/10.3745/JIPS.01.0024
24. Gupta S, Sinha S, Bhushan B (2020) The emergence of blockchain technology: fundamentals, working, and its various implementations. Work Various Implementations. https://doi.org/10.2139/ssrn.3569577
25. Kiayias A, Russell A, David B, Oliynykov R (2017) Ouroboros: a provably secure proof-of-stake blockchain protocol. In: Annual international cryptology conference, pp 357–388. https://doi.org/10.1007/978-3-319-63688-7_12

A Novel Image Encryption Technique for Grey and Color Digital Images

Shaik Kashif Hussain and S. Saheb Basha

Abstract Digital images require a high level of security while being sent over a transmission medium. If the data in these images is vulnerable to unauthorized use, this could lead to major problems. There are several techniques for protecting digital images. Encryption is the effective approach for protecting digital images. In encryption methods, confusion and diffusion are the two major stages. This paper introduces a novel chaotic map-based encryption technique for the encryption of digital images. Discrete wavelet transform (DWT)-based image blocks are the basis of a novel method for decomposing images is introduced. Then, the image blocks were permutated randomly, rotated, and arranged in a zigzag pattern to scramble the image blocks. The encrypted image is decrypted using a key generated from a chaotic logistic map. Through the use of security analysis, PSNR, and MSME, we are able to determine how effective our proposed method is at encrypting digital images. When compared to other methods, the results show that the proposed technique attain significant amount of performance.

Keywords Spatial domain · Image encryption · Chaotic system · PSNR · MSE

1 Introduction

The use of medical images to identify a wide range of diseases has become prominent due to the rapid advancements in medical equipment technology. As digital images are transmitted across many networks, preserving them has become a crucial issue in

S. K. Hussain (✉)
Research scholar, Department of Electronics and Communication Engineering, Jawaharlal Nehru Technological University Anantapur, Ananthapuramu, Andhra Pradesh, India
e-mail: kashif1919@gmail.com

S. Saheb Basha
Professor, Department of Electronics and Communication Engineering, G. Pulla Reddy Engineering College (Autonomous), Jawaharlal Nehru Technological University Anantapur, Ananthapuramu, Andhra Pradesh, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_66

these days. Security, reliability, and authenticity are required for the secure transfer of digital images. Using these images without authorization might compromise user's right to privacy. Additionally, if these images are vulnerable to even the slightest modification, it might lead to a misdiagnosis that endangers the lives of patients. In general, image cryptography [1, 2], image watermarking [3–5], and image encryption [6–8] can be utilized to secure digital images. Encryption is the quickest and most effective method of protecting image security which transforms a plain image into difficult to interpret form using a secret key. No one can restore the plain image without the secret key. The two major mechanisms of image encryption are diffusion and confusion. Digital image security research, including cryptography and image authentication, is consequently becoming more and more important. To minimize correlation and redundancy, numerous image encryption approaches [9–13] have been developed. In [14], Singh et al. presented an enhanced ElGamal encryption technique version which served as the foundation for the medical image encryption techniques. The increasing storage issue was resolved, and the long-term improvement was obtained. A novel medical image encryption method with random data insertion and pixel adaptive diffusion was presented out by Hua et al. [15]. In [16], a general optical encryption framework based on double random phase encoding (DRPE) was created for the purpose of encrypting medical images. An image encryption algorithm employing edge maps was developed by Cao et al. [17]. The algorithm is divided into three main components: permutation, bit-plane decomposition, and random sequence generation.

Digital images have a substantial association between nearby pixels; therefore, eliminating this correlation requires a permutation approach with a greater efficiency.

The four steps of the novel algorithm for encrypting images are image resizing, chaotic-based encryption algorithm, embedding image, and key generation. First, a new image splitting approach is employed to split the input image into blocks using discrete wavelet transform, and image resizing is done. Second chaotic-based encryption algorithm is used, and random permutation between blocks is used to alter the arrangement of the pixels inside the blocks and sub-blocks. Third, images are embedded after applying the algorithm. The logistic map, whose initial condition depends on the image, is employed in the fourth phase to construct the key. Finally, the secret key is used to change the image's pixel values. The rest of this work is arranged as follows. Section 2 examines in depth on the proposed methodology. The simulation analysis and results are presented in Sect. 3. Section 4 provides the paper's conclusion.

2 The Proposed Method

This section provides a detailed explanation of the main steps of the proposed approach for image encryption. The original image is converted into an unreadable image after being properly encrypted. Eventually, we use the decryption process to get the plain image.

2.1 Encryption

There are four stages in our technique for encrypting images. Image splitting and resizing is done in the first stage. In the second step, chaotic-based encryption is performed. In the third step, images are embedded after applying the algorithm. Key generation is presented in the fourth step using logistic map. Figure 1 shows an explanation of image encryption.

2.1.1 Image Resizing

The plain image is subdivided into non-overlapping blocks by using DWT. DWT decomposes the image into sub-bands based on high-pass and low-pass filtering operations. For a single level decomposition, it represents an image as four sub-bands, the first of which displays an approximation of the image in the low- and high-frequency components. Then, the process is followed by image resizing. The input image and the secret image are resized 256 × 256 pixels.

2.1.2 Applying the Algorithm

Using the chaotic sequences, the input image was rearranged to produce the permuted image and making the image completely unreadable. In the field of cryptography, chaotic maps are extremely important. These maps produce random numerical those are used in encryption as secret keys.

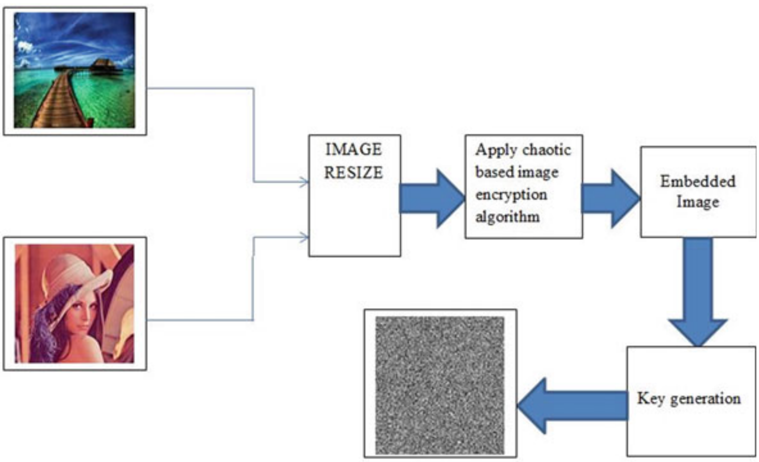


Fig. 1 Block diagram of image encryption

In the encryption process, chaotic maps assist in the confusion and diffusion procedures. Confusion and diffusion are applied repeatedly to each input to increase the level of scrambling. The process of confusion involves rearranging the image's pixel arrangement.

2.1.3 Embedded Image

Relatively high image quality is used in image embedding techniques such that the change between the original and embedded images cannot be seen with the bare eyes. Image encryption schemes scramble the image's pixels and reduce the correlation between them, making the encrypted image difficult to interpret.

2.1.4 Generation of Key

From a logistic map, the key used in the diffusion process is obtained. The logistic map is given by

$$X_{m+1} = b\ X_m(1 - X_m) \tag{1}$$

where b is the control parameter with range $0 < b \leq 4$, X_O is the initial value, and X_m is the output sequence with $0 < X_m < 1$.The map is chaotic when $b \in [3.75, 4]$.

These are the steps followed in the key generation:

1. To determine the beginning value of the logistic map, which is determined by the plain image P, use the equation below.

$$X_O = \frac{\sum_{k=1}^{P} \sum_{l=1}^{Q} M(k, l)}{P \times Q \times 255} \tag{2}$$

where P and Q indicate the plain image's number of rows and columns. Then, utilize the algorithm to determine the key.

$$\text{Key}(i) = \text{Key}(i) + \text{bin}(j \times i) \times 2 \tag{3}$$

where $i = 1$ to Q

$j = 1$ to 8.

2.2 Decryption

Utilizing the original key and reversing the encryption methods, we may retrieve the original image. We can encrypt or decrypt plain text into cipher text and then back into plain text with the use of a key and algorithm. The following is a description of the decryption process:

1. To obtain the scrambled image, do an exclusive OR operation on the encrypted image vector and key K.
2. Using the vector r, put each block back in its original location.
3. The inverse rotational operation is applied to sub-blocks.

3 Simulation Results

This part explains our algorithm's effectiveness at encrypting digital images. System with an Intel Core i3 2.27 GHz processor running MATLAB version 8.5.0.197613 is used to implement the approach. This method's analysis focuses on cover images and secret images that are nearly 256 by 256 in size.

Step 1: We consider two input images, i.e., plain image and the secret image for encryption, and the images are displayed as follows (Fig. 2).

Step 2: Image resize option is selected from pull down for both plain image and secret image, and resized images are displayed below (Fig. 3).

Step 3: After applying algorithm, the images are embedded and the embedded images are displayed below (Fig. 4).

Step 4: In the figure, after clicking encryption, it will display a dialog box where we need to set the secret key for encryption (Fig. 5).

Step 5: After generating the key for encryption, the encrypted image is displayed as below (Fig. 6).

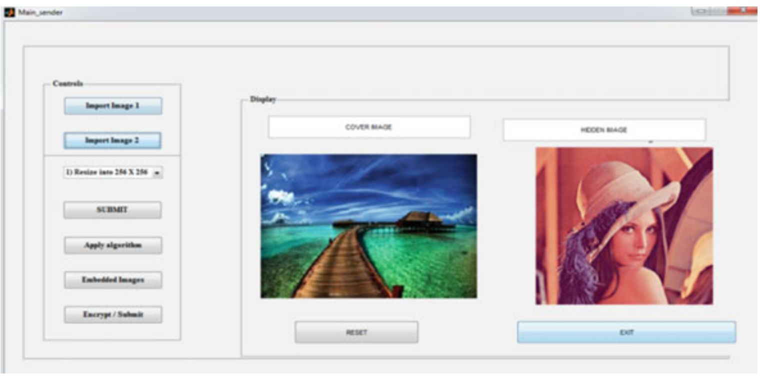


Fig. 2 Plane image and secrete image [18]

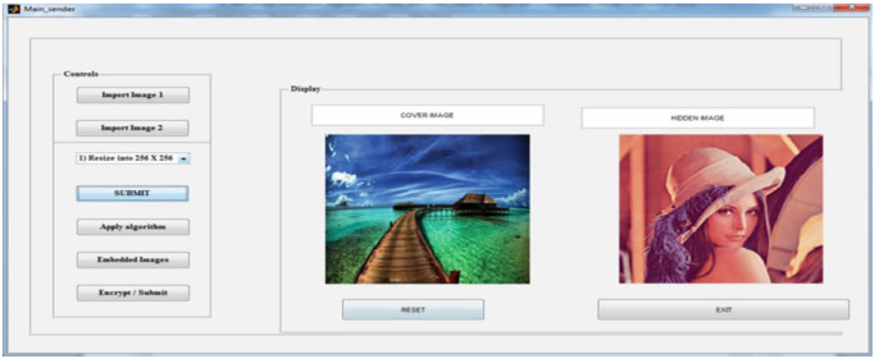


Fig. 3 Resized images [18]

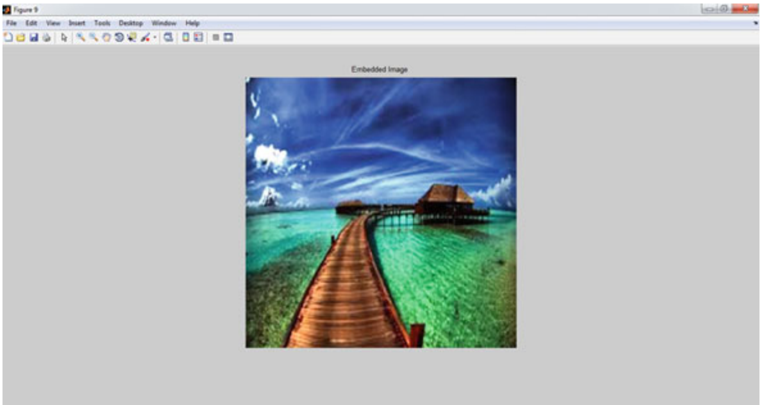


Fig. 4 Embedded image [18]

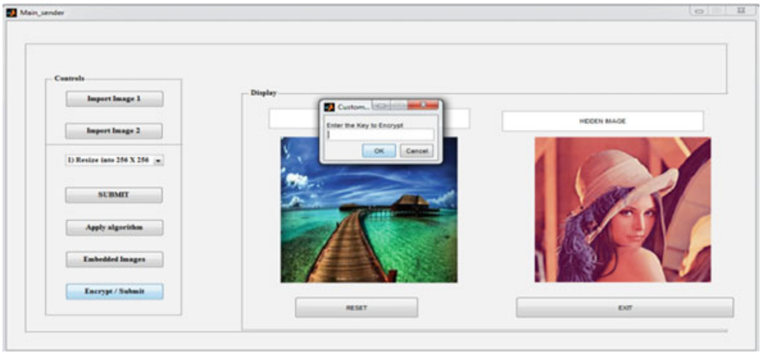


Fig. 5 Setting the secret key [18]

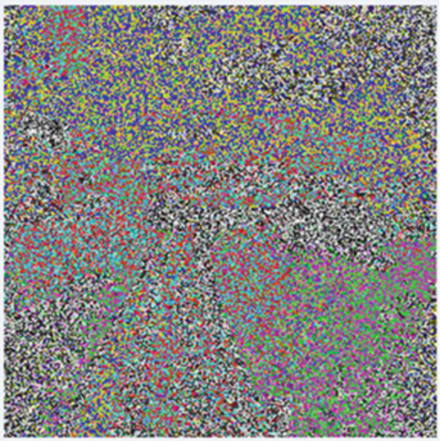


Fig. 6 Encrypted image [18]

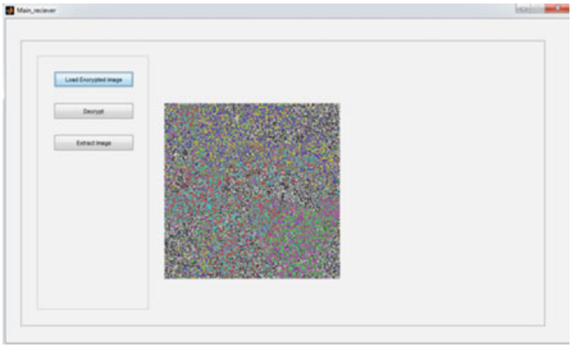


Fig. 7 Loaded encrypted image [18]

Step 6: In order to decrypt the encrypted image, we have to load the encrypted image tab as shown in below figure (Fig. 7).

Step 7: After loading the encrypted image, we have to enter the secret key which should be same as that of the key entered at encryption stage, and after using the secret key for decryption, the decrypted image is displayed (Fig. 8).

Step 8: After clicking the extract image tab, the hidden image is displayed as follows (Fig. 9).

3.1 Encryption Efficiency Analysis

The peak signal-to-noise ratio (PSNR), which analyzes the original and encrypted images, is used to assess the difference between them. The MSE is used to determine PSNR. The formulas for computing MSE, often known as the mean squared error, and PSNR are:

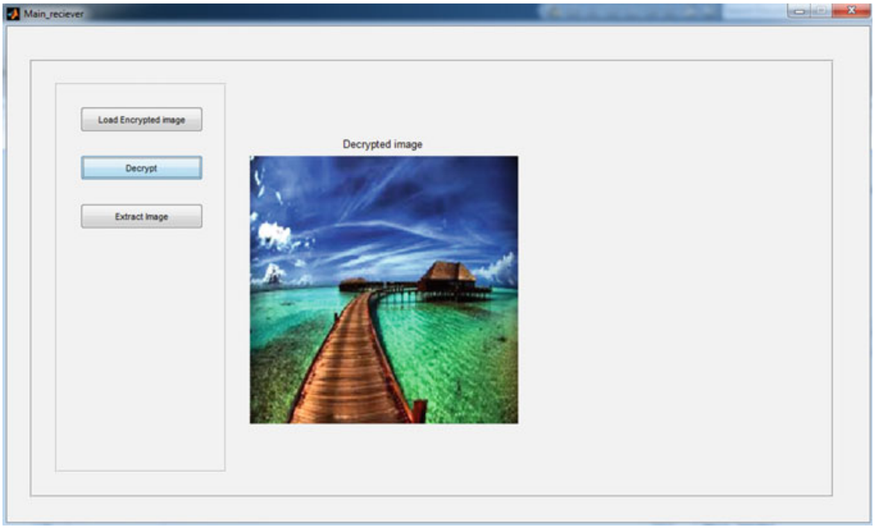


Fig. 8 Decrypted image [18]

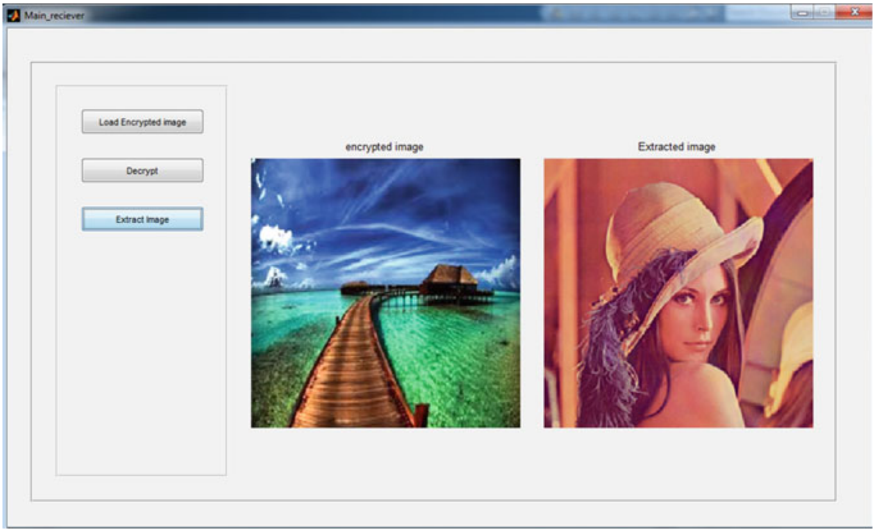


Fig. 9 Extracted image [18]

$$PSNR = 10 \times \log_{10}\left(\frac{255^2}{MSE}\right) \tag{4}$$

$$MSE = \frac{1}{PQ}\sum_{K=1}^{P}\sum_{J=1}^{Q}|O_r I(i, j) - E_n I(i.j)|^2 \tag{5}$$

Table 1 Analysis of our algorithm and other algorithms' PSNR and MSME values

Method	PSNR	MSE
Proposed method	31.4384	9.0557e+07
[7] sequence operations and chaotic systems	30.362	10.03276e+04
[8] A hybrid chaotic approach	30.00128	9.0691e+02

The $O_r I$ represents the original image, and $E_n I$ is used to represent the encrypted image. The PSNR numericals for various digital images are listed in the table. We can infer from the findings that our suggested technique is quite effective at encrypting digital images (Table 1).

4 Conclusion

This work presents a novel digital image encryption technique made up on image blocks and chaotic maps. Encryption is one of the effective techniques for protecting digital images. Using PSNR, differential attack, and MSE, the proposed algorithm's image encryption was evaluated. According to the results, the proposed algorithm is effective at encrypting digital images. The outcomes of comparing our approach to other recent encryption techniques show that the proposed approach has great potential for encryption of both gray and color digital images.

References

1. Liao X, Yin J, Guo S, Li X, Sangaiah AK (2018) Medical JPEG image steganography based on preserving inter-block dependencies. Comput Electr Eng 67:320–329
2. Usman MA, Usman MR (2018) Using image steganography for providing enhanced medical data security. In: Proceedings 15th IEEE annual consumer communication network conference (CCNC)
3. Hosny KM, Darwish MM, Li K, Salah A (2018) Parallel multi-core CPU and GPU for fast and robust medical image watermarking. IEEE Access 6:77212–77225
4. Hosny KM, Darwish MM (2018) Robust color image watermarking using invariant quaternion Legendre-Fourier moments. Multim Tools Appl 77(19):24727–24750
5. Hosny KM, Darwish MM (2019) Resilient color image watermarking using accurate quaternion radial substituted Chebyshev moments. ACM Trans Multim Comput Commun Appl 15(2):1–25
6. Vengadapurvaja AM, Nisha G, Aarthy R, Sasikaladevi N (2017) An efficient homomorphic medical image encryption algorithm for cloud storage security. Procedia Comput Sci 115:643–650
7. Liu J, Ma Y, Li S, Lian J, Zhang X (2018) A new simple chaotic system and its application in medical image encryption. Multim Tools Appl 77(17):22787–22808
8. Liu J, Tang S, Lian J, Ma Y, Zhang X (2019) A novel fourth order chaotic system and its algorithm for medical image encryption. Multidimensional Syst Signal Process 30(4):1637–1657

9. Shankar K, Elhoseny M, Chelvi ED, Lakshmanaprabu SK, Wu W (2018) An efficient optimal key based chaos function for medical image security. IEEE Access 6:77145–77154
10. Chen J, Chen L, Zhang LY, Zhu Z-L (2019) Medical image cipher using hierarchical diffusion and non-sequential encryption. Nonlinear Dyn 96(1):301–322
11. Ke G, Wang H, Zhou S, Zhang H (2019) Encryption of medical image with most significant bit and high capacity in piecewise linear chaos graphics. Measurement 135:385–391
12. Belazi A, Talha M, Kharbech S, Xiang W (2019) Novel medical image encryption scheme based on chaos and DNA encoding. IEEE Access 7:36667–36681
13. Mishra Z, Acharya B (2020) High throughput and low area architectures of secure IoT algorithm for medical image encryption. J Inf Secur Appl 53, Art no 102533
14. Laiphrakpam DS, Khumanthem MS (2017) Medical image encryption based on improved ElGamal encryption technique. Optik 147:88–102
15. Hua Z, Yi S, Zhou Y (2018) Medical image encryption using high-speed scrambling and pixel adaptive diffusion. Signal Process 144:134–144
16. Chen M, Ma G, Tang C, Lei Z (2020) Generalized optical encryption framework based on Shearlets for medical image. Opt Lasers Eng 128, Art no 106026
17. Cao W, Zhou Y, Xia L (2017) Medical image encryption using edge maps. Signal Process 132:96–109
18. Munson DC (1996) A note on Lena. In: IEEE Trans Image Process 5(1):3–3. https://doi.org/10.1109/TIP.1996.8100841

A Deep Learning Model to Detect People and Monitor Social Distance

Padmashree Desai, Saumyajit, C. Sujatha, S. M. Meena, M. Lalita, and Saurav Ansuman

Abstract Social distancing appears to be a very effective method of slowing the spread of contagious lung disease caused by influenza viruses that contaminate the nose, throat, and sometimes the respiratory after suffering of COVID-19 in the global battle. People are advised to limit their physical contact to lessen the virus's transmission hazard by physical or close contact. Artificial intelligence and deep learning are used to analyze real-time scenarios captured by CCTV cameras and determine whether crowd safety measures such as wearing masks or maintaining social distance among themselves have been followed. Our proposed work focuses on how a deep learning framework can detect social distancing in shared spaces and workrooms. The proposed human proximity detector system tracks whether individuals are preserving a safe distance between themselves or not by processing video streams captured in crowded scenes that are captured in public areas and workplaces. The proposed system can be integrated into surveillance systems in workplaces, factories, or shops to monitor whether employees keep a safe distance from one another at work. The YOLO deep learning model is adopted for object detection. A pairwise distance measuring algorithm is used for distance calculation between pairs of objects to decide whether they are maintaining a safe distance. The system determines whether or not an individual follows the social-distancing law and is classified as safe or unsafe. Both a live stream and a video feed validate the findings. Results indicated that the YOLO model performed well. YOLO v3 is recorded with the most effective results with a balanced frame per second (FPS) score of 23 compared to other methods. The accuracy of the proposed model YOLO v3 is 91.0%.

Keywords Social distance · Deep learning · Artificial intelligence · YOLO · FPS

P. Desai (✉) · Saumyajit · C. Sujatha · S. M. Meena · M. Lalita · S. Ansuman
KLE Technological University, Hubli, India
e-mail: padmashri@kletech.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_67

1 Introduction

Social distancing has gained a lot of importance during the battle of COVID-19. People started following minimum distances while walking on roads, in public places, and in discussions at offices. This motivated the proposed research work of detection of people on the road and finding out whether they are following safe distance or not. Study is based on the work carried out for the battlement of COVID-19 and computer vision and deep learning techniques applied for solving the spread of contagious respiratory disease caused by influenza viruses.

The coronavirus produces severe acute respiratory syndrome, which is an infectious condition. The sickness was initially found in December 2019 in Wuhan, China, and spread throughout whole world. When two people are close together, the virus travels largely between them, even though little droplets are produced by sneezing or coughing. Droplets that fall to the ground go in the air and into the human's body. Within the first three days, the infection is most contagious. Nausea, tiredness, and a dry cough are all frequent symptoms. The human consequences have been severe and negative, prompting a worldwide pause. Two instances of such symptoms include a headache and a sore throat. A real-time reverse transcription-polymerase chain reaction on a nasopharyngeal swab is the most common method of diagnosis (RTPCR). Based on symptoms and risk factors, chest CT imaging can also be used to diagnose persons at high risk of infection. To control the spread of this disease, the World Health Organization (WHO) coined the term "social distancing." Also, individuals have compulsorily used masks to protect themselves from COVID. Keep a two-meter distance between two persons to stay healthy. Following the COVID-19 epidemic, the Centers for Disease Control and Prevention (CDC) described social distancing as staying out of crowded places, avoiding public gatherings, and keeping a six-foot or two-meter gap between everyone, where practicable. As a result, adhering to the social-distancing rule is required and advantageous for living a safer and healthier life.

Emerging technologies have the potential to aid in the monitoring of social distancing. A current study [1] found that recent technologies such as networking, wireless communication, and AI permit or impose social distancing. The paper conversed potential basic theory, capacities, methods, and real-world scenarios for social distancing. The study in paper [2] categorized developing techniques into smart-space or human-centric categories and performed a SWOT analysis on the methods discussed. The paper [3] proposed a definite social-distancing monitoring style that detects and tracks pedestrians using YOLOv3 and deep sort, followed by computing a destruction value for non-social-distancing behaviors. Monitoring social distancing is also called as visual social-distancing (VSD) problem [4]. Different models based on sensing and machine learning technologies have recently been designed to aid in monitoring social distancing. Landing AI [5] proposed a model for detecting social distance that highlights people who are not following a physical distance of two meters which is prescribed by observing video from a surveillance camera footprint. In manufacturing plant also, alike system [6] was used to watch user movements and show alert messages. In addition to security cameras, LiDAR-based [7], and stereo camera-based [8] work was proposed illustrating that sensors other than surveillance cameras can also be useful.

Paper [9] discusses how combining a DNN model with the inverse perspective mapping (IPM) adapted method and the SORT for tracking resulted in a reliable system for detecting people and monitoring the social distancing. The model was trained on two complete datasets available during the study: Google Open Image and Microsoft Common Objects in Context (MS COCO).

Authors in [10] found the co-relation between the region's economic situation and the degree of social distancing. Several countries have adopted technological systems to combat loss due to pandemic [3]. Many developed nations use GPS technology to track the activities of contaminated and suspicious persons.

Authors in paper [1] present a survey of various evolving skills, with Wi-Fi, GPS, Bluetooth, computer vision positioning (localization), smartphones, and AI learning techniques that can play an essential part in various real-world social-distancing situations. Authors in [3] proposed a framework that detects humans using the YOLOv3 model and tracks them with localized bounding rectangles and assigned identification number (ID).

In [11], researchers created a drone-based automated model for monitoring social distance. The embedded camera in the drone and the YOLOv3 model assist in determining social distance and monitoring the front and side views of people wearing masks. In [12], the authors demonstrated the detection of human beings in a packed atmosphere. The approach is envisioned for people not following a social-distance limit of six feet amid them. In [10], the authors explored the association between the economic condition of the regions and the strictness of social distancing. Several nations have adopted knowledge-based solutions to address pandemic loss [3].

Our research aims to decide whether or not an individual follows the social-distancing law. The proposed work extracts frames from the video stream. It uses YOLO deep learning frame for object detection and distance measuring algorithm between pairs of objects to decide whether they are maintaining a safe distance. Social distance is measured between two frames of people from the centroids. People are categorized as safe or unsafe. A live stream and YouTube video feeds are used to validate the findings. It is also an exact number of social-distancing violations in a particular place. Detect human beings in the video feed, calculate the pairwise distance between two identified people concerning their centroid and based on n pixels apart, and label them safe or unsafe. Automate the surveillance for target users; social distancing in crowded places can be monitored without actually being present in the area. The system also provides ease of access for the people watching. Authors in [13, 14] explain the objection detection and tracking using deep leaning models. Authors in [15–20] discuss different solutions for following social distancing in a crowd.

2 Methodology

The proposed system's pipeline is shown in Fig. 1, which explains the processing of video frame, detection of people and computing pairwise distance, checking whether n pixels are apart, and marking them as safe or not.

2.1 *Three Stage Model*

The proposed model consists of three stages: detecting people, estimating inter-distance, and labeling people as safe or unsafe as a total solution for human proximity detectors. Figure 2 explains the different process of these three stages. The pre-trained YOLOv3 model on the COCO data provides a transfer learning technique. Transfer learning enhances detection performance, and a new layer of overhead training modifies the current architecture to obtain individual detection. In subsequent detection, centroid distance is computed for each bounding box using centroid information. The Euclidean distance formula measures the distance between each detected person's bounding box. After the computation of centroid size, the threshold value is used to determine whether the distance between any two bounding box centroids is within range (less than or equal to the pixels constituted). The information of the bounding box is stored in a vector representing a violation of pandemic rule, and the color of bounding box is changed/updated to red. An algorithm for centroid tracking is used for monitoring, which helps to identify the persons who interrupt or break the threshold of social distancing. The performance of the model reveals the count of social-distancing breaches and the people bounding boxes and centroids observed.

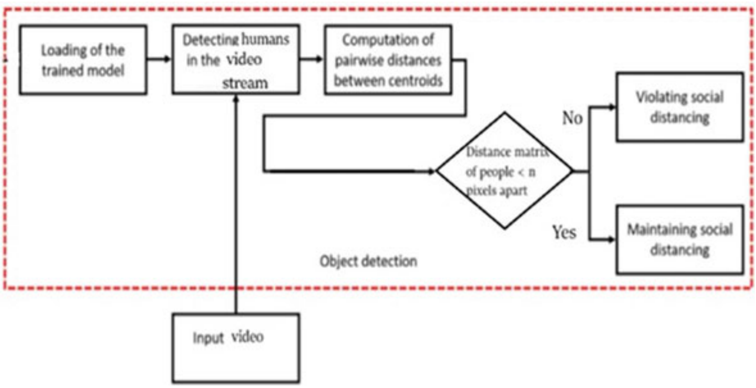


Fig. 1 Pipeline of the proposed system

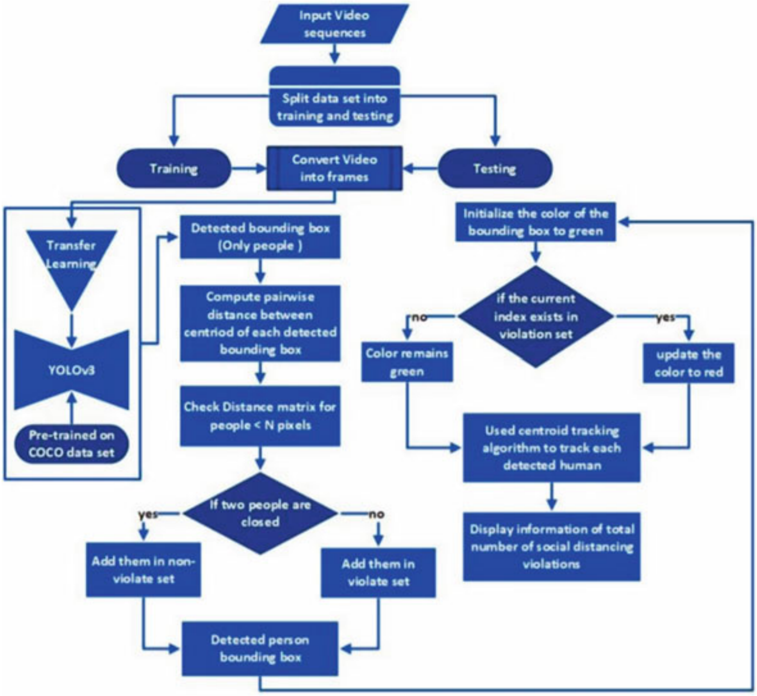


Fig. 2 System flowchart

2.2 Object Detection

In the system, the YOLO v3 trained model is used as transferred learning that can detect humans (people) in various situations, including variations in clothing, postures, lengths, with and without occlusion, and lighting conditions.

2.3 YOLO (You Only Look Once) Model

The YOLO detection model is shown in Fig. 3. The single shot detector (SSD) system is similar to YOLO. The approach solves the recognition of object as a problem of regression by considering a given image or video stream as input and calculating the coordinates of bounding box and the corresponding class label probabilities. YOLO has three tuning parameters: input size of network, the network for feature extraction, and anchored box.

To improve the accuracy, data augmentation was performed by randomly transforming the data while training. Using transform augmentation during training allows

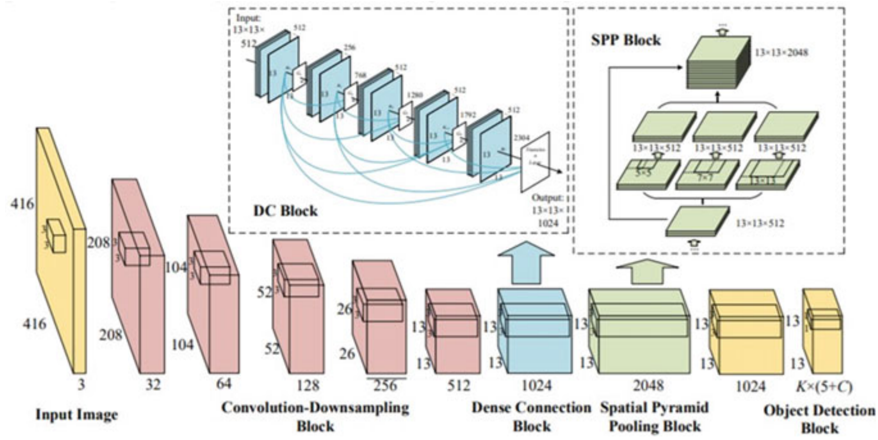


Fig. 3 YOLO architecture

for the random retention of images. The associated box labels are also horizontally flipped.

2.4 Centroid-Based Object Tracking

We have proposed the centroid-based object tracking tool to update the distance between two people continuously. The process of generating a unique ID for each initial collection of object detections (such as an input set of bounding box coordinates) is known as object tracking. Following that, it tracks each object as they pass through frames in a video while retaining its specific IDs.

Each detected object in each frame is given a set of the bounding box (x, y)-coordinates, according to the centroid-tracking algorithm. Any object detector can generate these bounding boxes as long as they are computed for each video frame. Any object detector can generate these bounding boxes as long as they are calculated for each video frame. We calculate the "centroid," or the central (x, y)-coordinates of the bounding box, and unique IDs are assigned.

The above step of calculating object centroids for next succeeding frame in the video stream is repeated. Distance calculation is done using Euclidian distance formula which represents the square root of the number of the squared differences between the two vectors, as given in Eq. 1 which is used to calculate the relative distance.

$$\text{Euclidian distance} = d(p, q) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \tag{1}$$

p, q two points in Euclidean n-space
n No of objects
q_i, p_i Euclidean vectors, starting from the origin of the space (initial point).

The main assumption in the algorithm for centroid tracking is that even if an object moves between frames, the distance between frames $F(t)$ and $F(t+1)$ concerning centroids will remain smaller compared to all other distances among objects of interest. The proposed centroid tracker algorithm associates centroids with the shortest Euclidean distance between objects. Assign a unique new object ID, store the bounding box centroid coordinates related to the object, and add it to the list of tracked objects. The pipeline of these steps is repeated for each frame for the input video stream. If an object is lost, disappears, or leaves the field of view, deregister the old objects for a total of N frames if they cannot be compared to any existing objects. When people are n pixels apart, the bounding box is green, implying they are safe, and if people are not n pixels apart, the bounding box is of red, stating the pair as unsafe.

3 Results and Discussion

Dataset

COCO dataset includes object identification, segmentation, and captioning on a wide scale. The name of the dataset, Common Objects in Context (COCO), literally means "everyday objects captured from everyday scenes." The dataset also provides the objects captured in the scenes with some "meaning." The dataset consists of 328 K images. The input video from the real world was used as the test data. Additionally, the 2017 release contains a new unannotated dataset of 123 K images.

The test video sequences are used to validate the model. Various video sequences are used to evaluate the testing results. The sample images show the model detects individuals in various scenic locations. Figure 4 shows the loss vs. iterations graph while training our model. Avg loss for training is 52% for 1800 iterations.

Sample frames are shown in Fig. 5 which indicates that there is no social-distance violation since all people are marked with green rectangle boxes by the automated framework. While in the sample frame shown in Fig. 6, the violation is detected; however, the number of people present in the scene is more as compared to Fig. 5, where all people are maintaining social distance, and therefore not a single violation is observed. Multiple people moving in the scene are detected and monitored in Fig. 7a, b. If people are too close, the framework detects the breach of social distance and marks the bounding box as red rectangles.

Table 1 compares proposed YOLO v3 with state-of-the-art methods referred for balanced frame per second (FPS) score and accuracy. The YOLO v3 model performed better than other models.

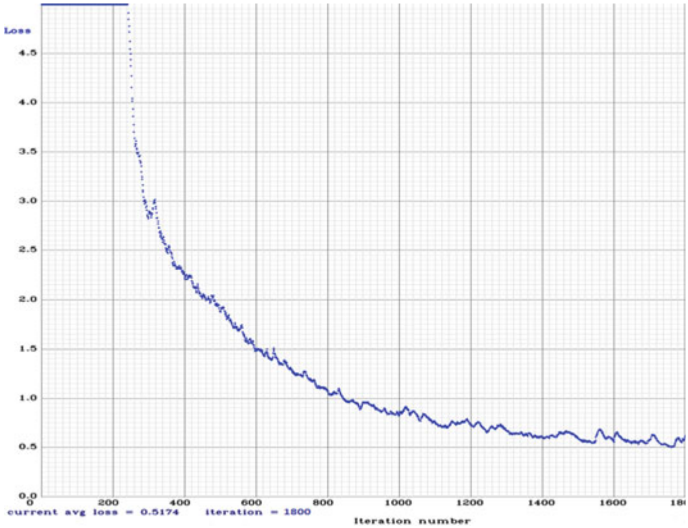


Fig. 4 Loss versus iterations graph

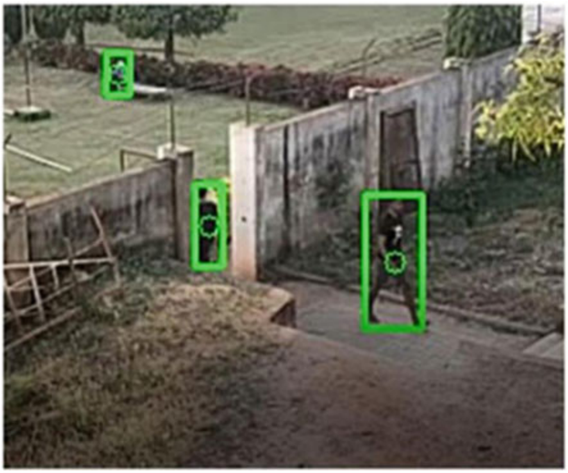


Fig. 5 No violation of social distance

4 Conclusion

The proposed system is a real-time deep learning-based framework for automating the process of monitoring social distancing through object detection and tracking approaches in crowded scenes captured through the CCTV. Each person is detected in real time using bounding boxes. The developed bounding boxes aid in identifying clusters or groups of people who satisfy the closeness property of the pairwise vectorized approach. The number of violations is determined by multiplying the number

Fig. 6 Three violations of social distance

a b

Fig. 7 **a, b** Multiple violations of social distance

Table 1 Comparison of YOLO v3 with state-of-the-art methods

Model	Accuracy (%)	FPS
Yadav [10]	91	10
Sener and Ikizler-Cinbis [21]	93.3	10
Liu et al. [22] (SSD300)	74.3	10
Liu et al. [22] (SSD512)	76.8	12
ResNet-50	86.5	13
ResNet-18	85.3	11
YOLO v3	91.0	23

of people crammed with the violation index term. The YOLO v3 demonstrated efficient performance with a balanced FPS of 23 and an accuracy of 91.0%. This method can be fine-tuned to better align with the corresponding field of view because it is susceptible to the camera's spatial position. Future scope includes use of better object detection models and distance measures.

References

1. Nguyen CT, Saputra YM, Van Huynh N, Nguyen N-T, Khoa TV, Tuan BM, Nguyen DN, Hoang DT, Vu TX, Dutkiewicz E et al (2020) Enabling and emerging technologies for social distancing: a comprehensive survey. arXiv preprint arXiv:2005.02816
2. Agarwal S, Punn NS, Sonbhadra SK, Nagabhushan P, Pandian K, Saxena P (2020) Unleashing the power of disruptive and emerging technologies amid covid 2019: a detailed review. arXiv preprint arXiv:2005.11507
3. Punn NS, Sonbhadra SK, Agarwal S (2020) Monitoring covid-19 social distancing with person detection and tracking via fine-tuned yolo v3 and deep sort techniques. arXiv preprint arXiv:2005.01385
4. Cristani M, Del Bue A, Murino V, Setti F, Vinciarelli A (2020) The visual social distancing problem. arXiv preprint arXiv:2005.04813
5. Fong MW, Gao H, Wong JY, Xiao J, Shiu EY, Ryu S, Cowling BJ (2020) Nonpharmaceutical measures for pandemic influenza in nonhealthcare settingssocial distancing measures
6. Alto P (2020) Landing AI named an April 2020 cool vendor in the Gartner cool vendors in AI core technologies. https://www.yahoo.com/lifestyle/landing-ai-named-april2020-152100532.html [Online; accessed 21 Dec 2020]
7. O. website of Indian Government (2020) Distribution of the novel coronavirus-infected pneumoni Aarogya Setu Mobile App. https://www.mygov.in/aarogya-setu-app/
8. Harvey J, LaPlace A (2019) Megapixels.cc: origins, ethics, and privacy implications of publicly available face recognition image datasets [Online]. Available: https://megapixels.cc/
9. Rezaei M, Azarmi M (2020) Deepsocial: Social distancing monitoring and infection risk assessment in covid-19 pandemic. Appl Sci 10(21):7514
10. Yadav S (2020) Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence. Int J Res Appl Sci Eng Technol 8(7):1368–1375
11. Ramadass L, Arunachalam S, Sagayasree Z (2020) Applying deep learning algorithm to maintain social distance in public place through drone technology. Int J Pervasive Comput Commun
12. Nadikattu RR, Mohammad SM, Whig P (2020) Novel economical social distancing smart device for covid-19. Int J Electr Eng Technol (IJEET) 8
13. Bagga NS, Desai P (2022) Person re-identification in multi-camera environment. In: 2022 2nd international conference on intelligent technologies (CONIT), pp 1–6. https://doi.org/10.1109/CONIT55038.2022.9847820
14. Kumar P, SC, Desai P (2022) Person tracking with re-identification in multi-camera setup: a distributed approach. In: 2022 2nd international conference on intelligent technologies (CONIT), pp 1–5. https://doi.org/10.1109/CONIT55038.2022.9848194
15. Courtemanche C, Garuccio J, Le A, Pinkston J, Yelowitz A (2020) Strong social distancing measures in the united states reduced the covid-19 growth rate: study evaluates the impact of social distancing measures on the growth rate of confirmed covid-19 cases across the united states. Health Affairs, pp 10–1377
16. Sulman N, Sanocki T, Goldgof D, Kasturi R (2008) How effective is human video surveillance performance? In: 2008 19th international conference on pattern recognition. IEEE, pp 1–3
17. Wang X (2013) Intelligent multi-camera video surveillance: a review. Pattern Recogn Lett 34(1):3–19
18. Gawde GS, Kalbande D (2020) Digital solution for enforcing social distancing. Available at SSRN 3614898
19. Feng Y, Marchal T, Sperry T, Yi H (2020) Influence of wind and relative humidity on the social distancing effectiveness to prevent covid-19 airborne transmission: a numerical study. J Aerosol Sci 105585:12
20. Venkateswaran J, Damani O (2020) Effectiveness of testing, tracing, social distancing and hygiene in tackling covid-19 in India: a system dynamics model. arXiv preprint arXiv:2004.08859

21. Sener F, Ikizler-Cinbis N (2015) Two-person interaction recognition via spatial multiple instance embedding. J Vis Commun Image Represent 32(63):3
22. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) European conference on computer vision. Springer, pp 21–37

An Empirical Review on Secure Edge Computing Architecture

Archana Kollu, Krishna Keerthi Chennam, and Deepa Mahajan

Abstract Edge computing offers some unique advantages over traditional models, where computing power is centralized at an on-premise data center. The method promotes processing of data on-site or in use while collecting Internet of Things (IoT) data at the edge rather than transmitting it back to the cloud. Endpoint data is gathered and evaluated to proceed with the subsequent actions. Edge computing provides high performance, availability, low cost, and high privacy for better security aspects in IoT. The cloud computing features extended to edge computing handle interruptions of sensitive applications. By shifting computationally intensive workloads to edge servers, the quality of computing, including energy usage and network transmission time, was improved considerably. Security features in edge computing emphasizes on secure data transfers that move through devices outside of centralized data centers. Edge computing is used in many applications where IoT device generated data can be easily processed and used.

Keywords Cloud computing · Edge computing · Internet of Things · Framework · Hardware · Privacy requirements and software

1 Introduction

Current advances in software and hardware methods, the lot of Internet of Things (IoT) devices like smart cameras, smart phones, smart vehicles, etc., have expressively improved, and for that reason, IoT applications and their devices have developed universal in modern digital society. The current developments from authorized users in edge cloud minimize edge processing computations. The use of the edge cloud for time-sensitive calculations, for cognitive services and IoT applications, is

A. Kollu (✉) · D. Mahajan
Pimpri Chinchwad College of Engineering and Research, Pune, Maharashtra, India
e-mail: archana.kollu@pccoer.in

K. K. Chennam · D. Mahajan
Vasavi College of Engineering, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_68

growing significantly. The services are to be provided on demand with in specified timeline fulfilling all application requirements [1]. But, most of the IoT strategies consume partial assets for storage, communication, and computing [2]. Even though they operate in dynamic edge servers, volatile edge computing, and highly distributed environments, it is not necessary to assume they are trustworthy or reliable because they cannot be used in the cloud servers as discussed in [2]. Several IoT strategies are able to upload reliable IoT data to edge servers to minimize the computational problems as in [3]. Various applications are more resource intensive and latency sensitive in edge cloud. The often and high unfindable latency between the end-users and cloud is version of the conventional cloud computing example in appropriate [2].

However, there may be a significant network transmission delay when directly uploading image-related data to cloud servers. To address the issue of network transmission delay, mobile edge computing can be used [3]. To more enhance the number of computing incomes available to end users, it is able to accompaniment mobile edge computing resources with the computing abilities of other end-users' systems, also acknowledged as fog computing [4]. Cloud computing is an Internet-based computing where the communal sources such as computing and storage facilities, data, applications, and software are retrieved and utilized on demand in a suitable pay as you go method [5]. However, the edge computing consumes the great possible to release the load on essential networks, its main block is the limited communication and computation abilities as associated with the cloud computing [2]. Though, IoT strategies from high latency and low bandwidth are feasible when collaborating with cloud servers [3]. This paper shows analysis about cloud-edge computing. Moreover, this paper specifies the major pros and cons in using edge computing.

The main objective of this paper is to increase edge enhancement to the computing resources, placed at the edge of the network. The problem of security enhancement minimizes the quantity of the data that need to be protected in the data center. To provide safety anxieties at individually contained point of the network edge, edge computing is employed. Large amount of data is highly susceptible to breaches, due to the similar security capabilities and built-in authentication. Edge computing methods are used to resolve the issues of high cost and storage space. But there is a chance of perilous data that might be misunderstood and demolished accidentally by an edge system.

The overall organization of the paper is given as follows: the overview about the secure edge computing discussed in Sect. 2. The literature survey and discussion about the existing method with its pros and cons are discussed in Sect. 3. Finally, the conclusion and future work is given in Sect. 4.

2 Overview of Secure Edge Computing

Edge computing is a developing method like cloud in which abilities is prolonged to the network edge. It expressively provides unified combination and minimalists the potential experienced in the framework of cloud computing with different application service vendors and providers [4]. Edge computing considers different various methodologies to make the network announcing possible to various attacks such as distributed denial of service (DDOS) attacks, middle attacks, wireless jamming, and denial-of-service (DoS) attacks. To address the edge computing issues, the coded computing framework was used [5]. The secure edge computing architecture diagram is shown in Fig. 1. The coded-edge computing policies are explained in the below sections.

2.1 Lagrange Code Computing

This computing is also called as linear coding system, and it used for data storage of edge campaigns. This is providing security concurrently privacy and distributed computing are resilient. The optimal retrieval edge utilizing Lagrange coding method is expressed in Eq. (1).

$$Y^t = (k - 1)\deg(f_t) + 1 \tag{1}$$

where Y^t is the threshold of optimal recovery, f_t Computation task

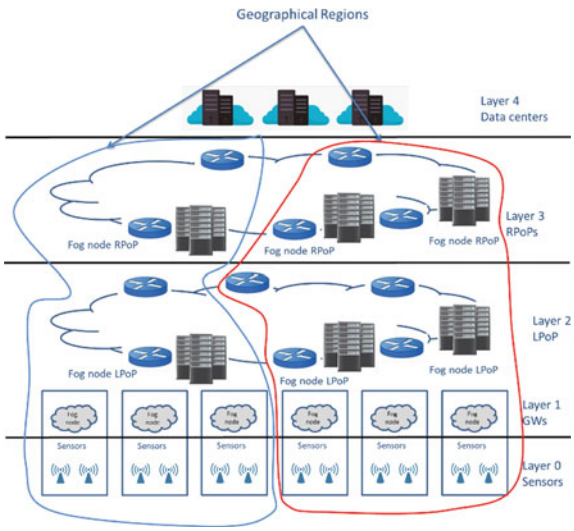


Fig. 1 Secure edge computing architecture

The encoding and decoding in Lagrange coding method are in polynomial evaluation and interpolation that can be done proficiently.

2.2 Contextual Combinatorial Multi Armed Bandit (CC-MAB) for Coded-Edge Computing

The coded-edge computing framework coding method is to be used for data encoding. Here, the encoding method is only performed one time for dataset [6]. After data encoding of Lagrange coding, the computation of every individual user and dimensions of input facts do not change.

The CC-MAB issues for coded-edge computing are expressed in Eq. (2).

$$\mathrm{Max}_{\{A^t\}_{t=1}^T} \quad \sum_{t=1}^{T} u\left(\mu^t, A^t\right) \tag{2}$$

2.3 Optimal Offline Policy

The success probability of every individual edge system is related to user. In common, the manipulator may not have success prospects of edge systems because of the indecision of the setting of edge network. The presentation of this feasible optimal policy is estimated by comparing with respect to this policy. This damage is named the regret of the policy which is properly expressed in Eqs. (3) and (4).

$$R(T) = E\left[\sum_{t=1}^{T} r\left(A^{t*}\right) - r\left(A^t\right)\right] \tag{3}$$

$$\sum_{t=1}^{T} u\left(\mu^t, A^{t*}\right) - u\left(\mu^t, A^t\right) \tag{4}$$

Contextual combinatorial multi-armed bandit (CC-MAB) issues allow the user to study the success probabilities of edge system over time by detecting the service superiority of every individual particular edge system and then make offloading verdicts adaptively.

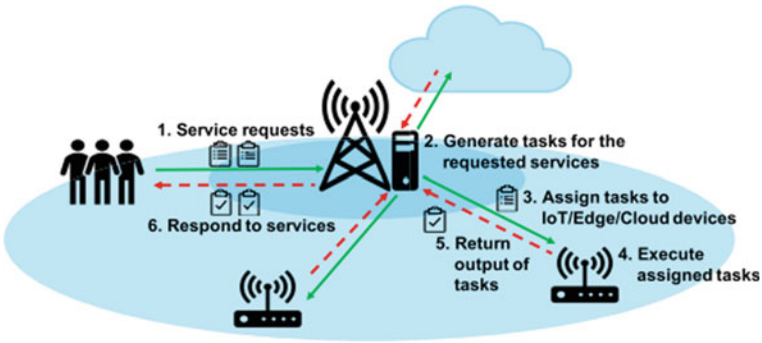


Fig. 2 Working mechanism of IoT-based edge computing

2.4 Online Coded-Edge Computing Policy

This policy is a monotonically and deterministic increasing function that can be used to predict the under-explored setting. This policy is accomplished as follows: hypercube identification phase, exploitation phase, initialization phase, exploration phase, and update phase. The IoT-based edge computing working mechanism is shown in Fig. 2. The IoT-based edge cloud computing mechanism works initially by service requests sent to devices, secondly devices generate tasks for the requested services, thirdly the task is assigned to IoT or edge or cloud devices, fourthly, the tasks will get executed, and in last returns the outputs and gives response to services.

3 Literature Survey

A brief evaluation of some contributions to the existing literatures is given below as summary of literature survey with respect to various computational methods, advantages, limitations, and performance measures discussed by various authors.

Summary of existing literatures

Author	Methodology	Advantage	Limitation	Performance measure
Ren et al. [5]	Collaboration of edge and cloud was used to increase the edge cloud efficiency	Reduce the weighted sum delay of all systems by investigating joint communication	The exact transmission delay is difficult to compute	System delay, optimal task splitting ratio, and resource allocation

(continued)

(continued)

Author	Methodology	Advantage	Limitation	Performance measure
Goudarzi et al. [7]	The weight cost method is used to optimize the energy consumption and execution time of IoT campaigns	To overcome the failure in the execution of tasks in runtime, lightweight failure recovery method is used	The cloud servers cannot find the possible outcomes	Weighted cost, execution time, and energy consumption
Singh et al. [8]	Proposed multi-access edge computing by using software-defined perimeter (SDP)	The port scanning and CPU usage were used to validate the efficacy of SDP	Before trying to flood the cloud service, the network servers do not validate to the controller	System delay, optimal task splitting ratio, and resource allocation
Kim et al. [9]	IoT-based edge computing through collaborative task scheduling method	The edge collaborative task scheduling is proposed as supplementary scheduling method in time critical tasks	The method was not accurate to find the best schedule for task implementation	Throughput, system delay, optimal task splitting ratio, and resource allocation
Hu and Li [6]	The dynamic request scheduling optimization algorithm was used for IoT tenders	The resource scheduling and joint request offloading method was used	Not appropriate for realistic applications	Energy consumption, welfare, and response rare
Guo et al. [10]	Proposed the delay guaranteed workload allocation to solve the workload problems	Here, the delay-based workload allocation method was used to minimize the energy consumption	The transmission rate does not exceed the bandwidth path	Energy consumption, accuracy, and average delay
Wang et al. [11]	Method to manipulate edge platform and network to ensure trust assessment and minimize the resource consumption	In edge platform, the service parameter pattern was recognized to enhance the efficacy of IoT cloud networks	The real-time devices were not able to share multiple applications	Energy consumption, throughput, accuracy, and average delay
Cui et al. [12]	Non-dominated sorting genetic algorithm (NSGA)-II is proposed	The single-objective problem was reduced by providing varying weights	The energy consumption is high in mobile edge computing (MEC)	Latency time, energy consumption, and transmission power

(continued)

(continued)

Author	Methodology	Advantage	Limitation	Performance measure
Lin et al. [13]	A mathematical model in edge computing devices is proposed for resource allocation problem	The IoT devices exchange limited amount of data which causes upper bound problem in every time period was reduced	The power consumption is high, limited resources available, and time delay was reported	Average delay, average drift, and frequency
Jha et al. [14]	IoT Sim-edge simulator for device heterogeneity	Prototype for scalable stimulation, identify the bottleneck performance	It is more complex in real-world applications and more time consuming	Average time, memory consumption, average latency, and iteration count

4 Conclusion and Future Work

Currently, a novel computing model of edge computing has developed to achieve the computing resources placed at the edge of the network, e.g., edge servers used for part of user service or all processing services [15, 16]. The IoT is susceptible to security fears, particularly interior attack that commonly arises in the network communication layer and physical device layer. This research was used to solve the above problems, and an edge-based IoT-Cloud manner with a service template and trust evaluation mechanism was recognized. The edge computing is made at the Internet edge with number of storage nodes and computing such as mobile fog nodes, gateways, edge server, and routers, which are near to the fundamental network. This research is very useful to edge computing and future things and will help the readers to understand the review of edge computing and also motivate more meaningful works. The important key attributed by edge is performance related to latency from cache, migrating, and virtual concepts. This shows the different attribute concerns like security, application, and scalability. In future, various extended studies are required to reach in depth survey of edge cloud computing tactics and methods.

References

1. Yang CS, Pedarsani R, Avestimehr AS (2021) Edge computing in the dark: leveraging contextual-combinatorial bandit and coded computing. IEEE/ACM Trans Netw 29(3):1022–1031

2. Prasad Kantipudi MVV, Aluvalu R, Velamuri S (2022) An intelligent approach of intrusion detection in mobile crowd sourcing systems in the context of IoT based SMART city. Smart Sci. https://doi.org/10.1080/23080477.2022.2117889
3. Wang S, Ding C, Zhang N, Liu X, Zhou A, Cao J, Shen X (2019) A cloud-guided feature extraction approach for image retrieval in mobile edge computing. IEEE Trans Mob Comput 20(2):292–305
4. Chennam KK, Uma Maheshwari V, Aluvalu R (2022) Maintaining IoT healthcare records using cloud storage. In: Nath Sur S, Balas VE, Bhoi AK, Nayyar A (eds) IoT and IoE driven smart cities. EAI/Springer innovations in communication and computing. Springer, Cham. https://doi.org/10.1007/978-3-030-82715-1_10
5. Ren J, Yu G, He Y, Li GY (2019) Collaborative cloud and edge computing for latency minimization. IEEE Trans Veh Technol 68(5):5031–5044
6. Hu S, Li G (2019) Dynamic request scheduling optimization in mobile edge computing for IoT applications. IEEE Internet Things J 7(2):1426–1437
7. Goudarzi M, Wu H, Palaniswami M, Buyya R (2020) An application placement technique for concurrent IoT applications in edge and fog computing environments. IEEE Trans Mob Comput 20(4):1298–1311
8. Singh J, Bello Y, Hussein AR, Erbad A, Mohamed A (2020) Hierarchical security paradigm for iot multi access edge computing. IEEE Internet Things J 8(7):5794–5805
9. Kim Y, Song C, Han H, Jung H, Kang S (2020) Collaborative task scheduling for IoT-assisted edge computing. IEEE Access 8:216593–221660
10. Guo M, Li L, Guan Q (2019) Energy-efficient and delay-guaranteed workload allocation in IoT-edge-cloud computing systems. IEEE Access 7:78685–78697
11. Wang T, Zhang G, Liu A, Bhuiyan MZA, Jin Q (2018) A secure IoT service architecture with an efficient balance dynamic based on cloud and edge computing. IEEE Internet Things J 6(3):4831–4843
12. Cui L, Xu C, Yang S, Huang JZ, Li J, Wang X, Ming Z, Lu N (2018) Joint optimization of energy consumption and latency in mobile edge computing for Internet of Things. IEEE Internet Things J 6(3):4791–4803
13. Lin R, Zhou Z, Luo S, Xiao Y, Wang X, Wang S, Zukerman M (2020) Distributed optimization for computation offloading in edge computing. IEEE Trans Wireless Commun 19(12):8179–8194
14. Jha DN, Alwasel K, Alshoshan A, Huang X, Naha RK, Battula SK, Garg S, Puthal D, James P, Zomaya A, Dustdar S (2020) IoT Sim-edge: a simulation framework for modeling the behavior of Internet of Things and edge computing environments. Softw Pract Experience 50(6):844–867
15. Jin W, Xu R, You T, Hong YG, Kim D (2020) Secure edge computing management based on independent microservices providers for gateway-centric IoT networks. IEEE Access 8:187975–187990
16. Zhang P, Jiang C, Pang X, Qian Y (2020) STEC-IoT: a security tactic by virtualizing edge computing on IoT. IEEE Internet Things J 8(4):2459–2467

Rumour Detection in the Political Domain from Twitter Using Machine Learning Techniques

Syed Asadullah Hussaini, Meer Tauseef Ali, Mahamood Ali, and P. Vishvapathi

Abstract With the evolving improvements of technology, online social messaging systems have become an important means for communication of information. Nowadays, social messages are difficult to remove from the daily lives of the people due to its instant transferring capabilities. Even the renowned organizations are checked by people based on their uploaded social organizational profiles, which can be easily captured through social messaging sites. Political organizations started to efficiently make use of various social media platforms (e.g. Twitter, Facebook, LinkedIn, Snapchat and many more) to propaganda their development and welfare schemes to attract the voter's attention to win the elections. False promises and political rumours made by these selfish politicians has been a serious threat for this society. Several researches based on topic-independent political rumour detection on Twitter have been well-known to everyone. In this paper, we proposed a rumour detection algorithm that has the capability to monitor political rumour messages especially on Twitter. For this, we took a subset of characteristics, apart from Sample Stream and Filter Stream. The proposed system is checked by taking the real-time data set and found that nearly 75.2% of exclusive political rumours are traced. Exclusively, the third largest smaller scale blogging platforms, i.e. Twitter found to be grown drastically for the use of political rumours apart from WhatsApp and Facebook. Both, Sample Stream and Filter Stream functions simultaneous one after the other and play their vital role of transmitting of tweets with certain pre-set criteria. It involves three (3) filtering criteria's: (a) Specific terms, (b) Targeted Twitter user and specific tweets chosen for a particular located area. At this specific, we start indulging our proposed rumour detection algorithm that efficiently starts investigating. We also conclude that efficacy is increased by 99.83% with the use of an ANN model.

S. A. Hussaini
ISL Engineering College, Hyderabad, India

M. T. Ali · P. Vishvapathi
Deccan College of Engineering and Technology, Hyderabad, India
e-mail: vishvapathi@deccancollege.ac.in

M. Ali (✉)
Shadan College of Engineering and Technology, Hyderabad, India
e-mail: mahamood.mtec@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_69

Keywords Twitter social network · Rumour detection system · Sample Stream and Filter Stream · Artificial neural network (ANN)

1 Introduction

1.1 Background Knowledge

From the past years, social messages are tough to remove from the daily lives of the people due to its outperforming message transferring capabilities. Twitter outpaced the older news dissemination system. It is often observed that news is first broken in Twitter space and then from the electronic area. An individual's and even the renowned organizations are checked by people based on their uploaded social organizational profiles, which can be easily captured through messaging apps. Political organizations started to efficiently make use of various social media platforms (e.g. Twitter, Facebook, LinkedIn, Snapchat and many more) to propaganda their development and welfare schemes to attract the voter's attention to win the elections. The message transferring system and acknowledging system in transferring information among the innumerable population has gave a big impact in the history of communication system. The tactics and notifications of messaging system in shortcuts led the peoples life easier and safer by saving lot of time in conveying of useful information and news as well. At the same time, fake and lying information also is on top of the notch in social media, especially political news which is indirectly a disturbing etiquette of the society distracting the developments of the society which needs stringent action by the government in the secular aspects. The objective is to catch hold of untruthful political rumours microblogs which are transmitted through innumerable social messaging systems. Apart from this, catching such online social criminals who intentionally or unknowing send such fake political microblogs and inform the concerned security surveillance department.

1.2 Problem Statement

False rumours made by political goons through social media sends wrong impact that tries to influence the ethics of people that motivates to create nuisance that disturbs the harmony among the people of society. There must be spying eye that monitors the centralized messaging servers to control such political rumours using surveillance systems to check the authenticity of those microblogs at social networking sites. Especially, political rumour microblogs needs to be blocked or protected, before it reaches the innocent social community. To implement, these rumours microblogs required to be browsed intelligently in authorized political databases to know its genuineness by providing highest level of security [11]. If the political news is authentic then only the server grants permission to pass news to other social users, otherwise the servers must initiate alerts (or) pop-ups to security department intimating the facts of people who initiated disturbances in the society for monetary benefits or political favours. Blocking of fake tweets made through politicians is not so easy task, as it involves in-depth pre-processing of microblogs [1].

In Twitter, trillions of social users make use of tweets and re-tweets up to a maximum limit of 140 characters in a minute. This might generate excessive loads on the servers for disseminating of news from multiple sources to innumerable locations. Further, if we implement surveillance systems on such servers is not so easy task, it means excessive high-end processing servers needs to be installed. To demonstrate it, we have taken some political datasets and tested on our own servers to catch such fake political rumours [9]. Many of the existing contributions exist that identifies rumours from different domains restricted to one particular domain, whereas the political fake rumours will vary from time to time, location to location, topic to topic with different agendas. To overcome, we attempted by collecting the information pertaining to various political rumours and tried our best to tackle such rumours and proposed the political rumour detection strategy especially for Twitter [10]. They also filter the discussion structure by only evaluating the most prominent individuals or the network's most crowded conversations.

Anyhow, such rumour detection system shouldn't restrict itself to a particular subject of political fake category, but also should have an ability to catch hold of various categories of political microblogs especially in Twitter. The rumour state and its problem are represented as follows:

- **SU** (set of online user's) in social media $u_i \in \mathrm{SU}$, u_i is every user in chat session.
- **Rumour State**: (r, u_i, t) or $\boldsymbol{r^t}$ denotes rumour state of user u_i.

$$
\begin{aligned}
r_i^t \in \{0, 1\} \text{ if } r_i^t &= 1 \text{ is rumour and} \\
\text{if } r_i^t &= 0 \text{ is not a rumour}
\end{aligned}
$$

- **User attribute matrix:** Let $\mathbf{M^t}$ be a $|\mathbf{U}|$ x $|\mathbf{GA}|$ every user in **SU** has group of attribute GA
- **Edge set**: Let $\mathbf{E^t} \subseteq \mathbf{U_i} \times \mathbf{U_j} \times \mathbf{C}$ be a set of edges between user,

$$
e = \left(u_i, u_j, c\right) \in \mathrm{E^t}
$$

where $c = 0$ shows u_i following or is followed by u_j, $c = 1$ signals positive communication where $c = 2$ represents negative communication

- **Attribute augmented network**: $\mathrm{G^t} = (\mathrm{U^t}, \mathrm{E^t}, \mathrm{M^t}, \mathrm{R^t})$
- **Rumour State**: Given a series of **T** partially labelled

$$
\left\{\mathbf{G^t} = \left(\mathbf{U^t}, \mathbf{M^t}, \mathbf{R^t}\right)\middle|(\mathbf{t} = \{\mathbf{1}, \mathbf{2}, \ldots \mathbf{T}\})\right\}
$$

$\mathbf{f} = \{G_u^1, G_u^2, \ldots G_u^t\} \rightarrow \{R^1, \ldots R^T\}$ the function predicts the unlabelled rumour states of users.

The idea is to build a rumour detection system as depicted in the above function (f), from social media microblogs tweets sent by the users in the social network such that when rumour is found, it has to be predicted and notified [9].

2 Literature Survey

An openness will definitely affects those broad communications and social online life. There are political goons who need to take maximum benefits of this to regain their political goals and build their career through the use of fake political news. They intentionally spread misleading information in the form of news to manipulate people in numerous ways. There are a slew of websites dedicated only to disseminate misleading information, which is untruthful news, publicity materials, manufactures and fear-mongering ideas under the guise of real news. The primary goal of fake news websites is to sway public opinion on certain topics (*for the most part of political career*).This may be exclusively seen in Ukraine, the United States of America, the United Kingdom, Russia and a variety of other countries. As a result, fake news is a global problem and a huge challenge to deal with especially in politics. The survey carried out in political rumour domain is given in Table 1, that gives the summary of how political rumours presently exists and the technique devised to resolve is elaborated.

3 Proposed Rumour Detection Model with ANN

In Python "tweepy", module acts as a package that assists as a streaming API for Twitter. It gathers data as raw tweets. Usually, two (2) methods to gather tweets using API are: *Sample Stream* and *Filter Stream.* Sample Stream simply delivers small, arbitrary samples extracts tweets which have been stream lined in real time during the tweeting process. Filter Stream sends back only those tweets which meets pre-defined criterion. The abilities of filtering are based on three (3) criteria's as depicted below.

- Trace/find a particular term from the tweets.
- Finding a Twitter exclusively only by their genuine names (or) sometimes nick name.
- Sometimes tweets that arrives from a particular location (world-wide).

Data streaming in Twitter is a social microblog with millions of users from across the world, enabling for real-time information dissemination to a huge audience and allowing users to do various actions. It is commonly observed that news is first broken

Table 1 Political rumour models, strategy used and drawbacks

Sl. No.	Title	Objective	Strategy/Technique	Remarks
1	Exclusive active finding of political rumour identification in Twitters [1]	Not only Twitter, but also social apps brought for scrutinizing for lack of its improper filtering issues that lead to various crimes such as which lead to the broadcasting of trolling, oppression and other un-ethical behaviours	The rules-based technique is employed for checking false political microblogs from Twitter by checking extreme case of online social users	Drawback is it is exclusively depended on fixed parameters
2	COVID-19 health checking related rumour and misinformation in Twitters [2]	They constructed anew partial features by embedding potential as well as network related to COVID-19 health problems from specific domain for various types of diseases	Pipeline method of the rumour detection process	The accuracy of 89% when tested on real data sets is identified
3	Crossing of topic for rumour identification in health domain [3]	Rumour detection in microblogs at various stages by accumulation of tweets, where a little effort is applied to predict rumour at various stages by picking each of the tweet	Strength of the parameter set is taken from earlier works as a pipeline	Crossing of topic considers only one part as a test set and other part as a training set in the experiments, which may not be accurate
4	Programmed rumour prediction from microblogs [4]	Mine the available microblogs forwarded via various means of network	Machine learning technique is employed to predict such tweets	Failed in identifying longest texts and multimodal instant spim sent in social sites
5	Big Data and eminence data for fake and mediocre information [5]	The issues involved are gathering of qualitative data of untruthful and genuine fake news based on equalized distribution of topics	Natural language processing (NLP) uses the feature approach, which involves the extraction and analysis of linguistic cues for identification of specific target phenomena	In order to detect fake news, the method requires large amount of training data

(continued)

Table 1 (continued)

Sl. No.	Title	Objective	Strategy/Technique	Remarks
6	Aspect stage sentiments analysing using NLP and deep methods [6]	Sentiment is a NLP technique can be found in comment or tweet. By using sentiments of various users exclusively based on a particular aspect extraction is done by employing of deep learning methods	CNN used for image classification, and an NLP is used for processing of text, i.e. sentiment analysis	CNN method also sometime fails to identify valid aspects

in the Twitter domain, then in the electronic area. An individual may have a variety of social media profiles on various platforms that identify his digital presence, such as Twitter, Facebook and Snapchat. We were able to collect 1198 tweets using the extreme user keyword thanks to Twitter's REST API capabilities. We get the data through a Twitter API called tweepy, which returns a restful array with the topic and language you provided (Here we are processing only English). We divided the rumours into two categories: rumours and non-rumours [7].

Finally, we have gathered the data in *JSON* format, and we are ready to begin pre-processing. This model includes an effective pre-processing phase that removes meaningless symbols from tweets, removes URLs, hashtags, tokenization and stop-words and then returns meaning full data after removing stop words, usernames and other variables. We also organize data in the form of data attributes or terms like id, user id, username, text, creation date, retweet count, favourite count, texthash, assertion, topic, rumour, with 11 columns of attribute and 1198 rows of tweets. We'll collect the overall impression and use an artificial neural network to extract the topic in order to classify it, and we'll end up with a 7% rumour classification and a 93% non-rumour classification.

Sample Input Test data is chosen from Textual data and retweets. This test data is fed to the proposed algorithm which is shown in Fig. 1.

4 Experimental Analysis and Comparative Results

Earlier works after the developments of SVM, it was used broadly and successfully used in Simulated Annealing (SA) To overcome those limitations, our research suggests that ANN executions would result in a better arrangement by combining the best of faux neural system with fluffy logic.

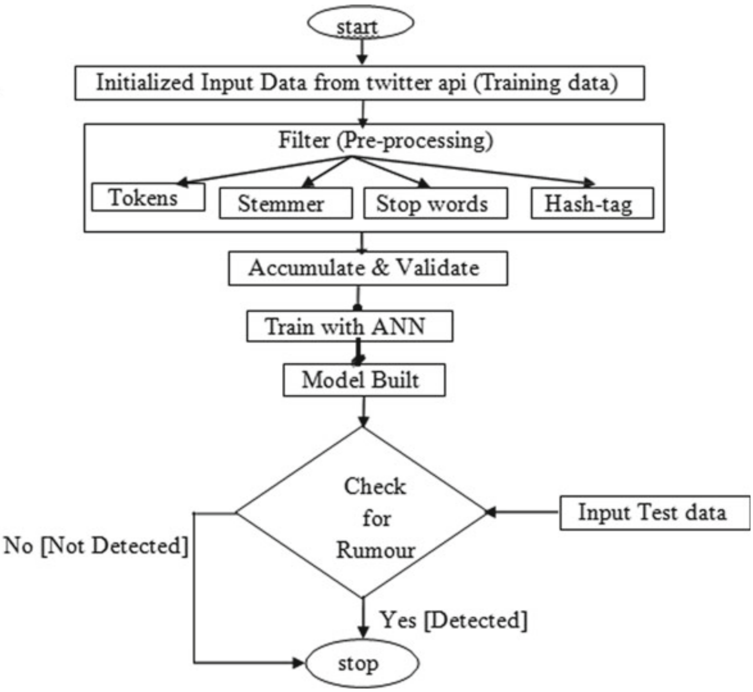


Fig. 1 Algorithm cum flow chart of proposed rumour detection algorithm using ANN

4.1 Pre-processing Stage

Pre-processing for Twitter microblogs which is a tedious task where the social users posts will contain mixed type of conversation which may become very tedious task to filter those post which has fake political rumours [8].

4.2 Results and Comparative Analysis

In supervised artificial intelligence algorithms, evaluation metrics are typically employed to measure the algorithm's efficacy with the aid of confusion matrix to evaluate rumours and non-rumours. The results obtained using our proposed ANN model are depicted in Fig. 2.

Comparative analysis is done using various algorithms, namely *random forest, decision tree, KNN* and proposed *ANN* are shown in Table 3.

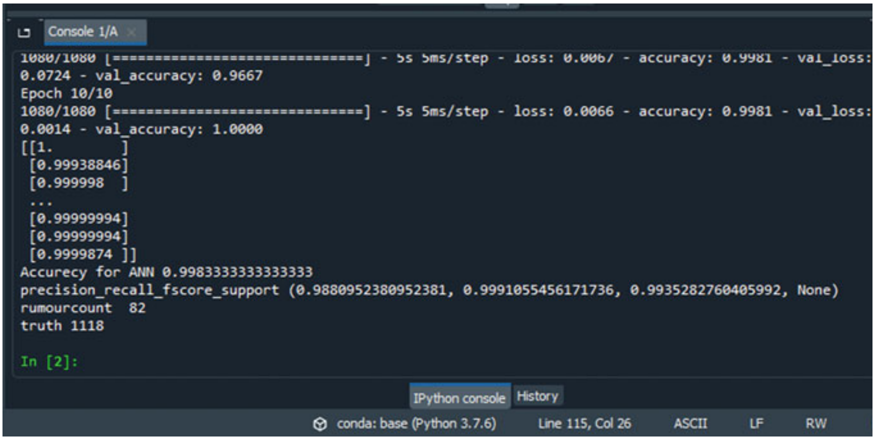


Fig. 2 Results obtained using proposed ANN model showing an accuracy of 99.83%

Table 3 Accuracy, precision, recall and F1 measure values using various algorithms

Methods used	Accuracy (%)	Precision (%)	Recall	F1 score
Random forest	0.94	0.93	0.65	0.76
Decision tree	0.90	0.89	0.25	0.39
KNN	0.93	0.93	0.48	0.63
Artificial neural network	0.99	0.98	0.99	0.98

5 Conclusion and Future Scope

On social networks, there has been a lot of research in the field of rumour detection. However, mediocre studies have seriously worked on microblogs such as Twitter. By demonstrating the spread of social media in all spheres of life, we highlight the necessity for new approaches to rumour identification that target rumours other than political rumours. Exclusively, the third largest smaller scale blogging platforms, i.e. Twitter found to be grown drastically for the use of political rumours apart from WhatsApp and Facebook. Both, Sample Stream and Filter Stream functions simultaneously used one after the other which plays their vital role of transmitting of tweets with certain pre-set criteria. It involves three (3) filtering criteria's: (a) Specific terms, (b) targeted Twitter user and (c) specific tweets chosen for a particular located area. At this specific point, we started indulging our proposed ANN rumour detection model that efficiently started investigating the tweets and showing the accuracy of 99.83%. We conclude that efficacy is increased with the use of ANN.

Future work could be directed towards two broad classes of networks by taking the behaviour aspects and psychological stressness from the tweets and other social networks [12, 13]. The extensive analysis of implicit rumour in political domain by these political goons may also leads towards unidentified suspicious activities that need to be focused from time to time by embedding surveillance cameras around their houses and whereabouts of their locations [14].

References

1. Parsa F, Mohammad MAP, Hamidian S, Diab M (2021) Active learning for rumor identification on social media. In: Findings of the association for computational linguistics: EMNLP. Association for Computational Linguistics, pp 4556–4565
2. Gautam KS, Dirkson A, Majchrzak TA (2021) An exploratory study of COVID-19 misinformation on Twitter. Online Soc Netw Media J 22:201–218 (Elsevier)
3. Ali G, Malik MSI (2022) Rumour identification on Twitter as a function of novel textual and language-context features. Springer, Multimedia Tools Application
4. Zongmin L, Qi Z, Xinyu D, Yanfang M, Shihang W (2021) Social media rumor refutation effectiveness: Evaluation, modelling and enhancement. Inform Process Manag 58(1) (Elsevier)
5. Thi Dieu Linh N, Lu Z (2021) Eminent role of machine learning in the healthcare data management. In: Data science and medical informatics in healthcare technologies. SpringerBriefs in Applied Sciences and Technology. Springer, Singapore
6. News Use Across Social Media Platforms in 2020. Pew research center (2021) (online). https://www.pewresearch.org/journalism/2021/01/12/news-use-across-social-media-platforms-in-2020/. Accessed 21 Nov 2021
7. Meng Z, Yu SR, Li, Jiang G, Song Y (2020) Dynamic features based rumor detection method. In: 2020 Chinese control and decision conference (CCDC). IEEE, pp 379–384
8. Pandey S, Gupta R, Kumar J (2022) Identification of fake news using machine learning techniques. In: Pundir AKS, Yadav N (eds) Recent trends in communication and intelligent systems. Algorithms for Intelligent Systems. Springer
9. Shu K, Deepak M, Suhang W, Lee D, Huan L (2020) FakeNewsNet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media. Big Data J 8(3)
10. Zubiaga A (2021) Fake news detection: a survey of evaluation datasets. PeerJ Comput Sci J Natl Libr Med
11. Mahmood MA, Qaseem MS, Rasheed MA, Khan MKA (2020) ESMD: enhanced suspicious message detection framework in instant messaging applications. In: Fourth international conference on inventive systems and control (ICISC), pp 777–784
12. Ali MM, Hajera S (2018) Psychological stress detection from social media data using a novel hybrid model. Int J Intell Syst Appl Eng 6(4):299–305
13. Ali MM, Noorain S, Qaseem MS, ur Rahman A (2023) Suspicious human behaviour detection focusing on campus sites. In: Chaurasia MA, Juang CF (eds) Emerging IT/ICT and AI technologies affecting society. Lecture Notes in Networks and Systems, vol 478. Springer
14. Juan C, Guo J, Li X, Jin Z, Han G, Jintao, Li (2018) Automatic rumor detection on microblogs: a survey. IEEE

Review on Recent Advances in Text Summarization Techniques

M. Vinitha and S. Vasundra

Abstract Recently, there is a massive increase in the volume of text information from different kinds of resources. The volume of text is a valuable resource of information and knowledge that must be efficiently summarized and made accessible. Document summarization will provide a tool for rapidly understanding text document collections and has numerous real-time applications. The two efficient approaches of semantic similarity and clustering are used to provide an effective overview of large text sets. Summarizing a large amount of text is time-consuming and challenging, specifically, while the computation of semantic similarity is considered in process of summarization. Summarization of text which is collected includes intensive processing of text and computations to produce the summary. In this study, the various summarization processes and the shortcomings and effectiveness of the several methods are also described.

Keywords Textmining · Rouge · Automatic text summarization · Summarizer · Extractive summarization · Abstractive summarization · Clustering · PGAN-ATSMT · Semantic similarity

1 Introduction

In text mining, the most challenging and significant problem is text summarization. It provides numerous advantages to users, and numerous productive real-time applications are also developed by utilizing text summarization [1]. When text is summarized, a huge number of text documents are gathered and transformed into a compact and reduced text document that represents the summary of the original text collections [2, 3]. The document which is summarized will assist to understand the

---

M. Vinitha (✉) · S. Vasundra
Department of CSE, JNTUA, Ananthapuramu, India
e-mail: vinitha@rguktong.ac.in

S. Vasundra
e-mail: vasundras.cse@jntua.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_70

gist of huge text collections and also save more time by ignoring reading of every separate document in a huge text collection [4, 5]. Mathematically, summarization of text is a function of transforming large information of text into small information of text in a way that the small information of text will carry the total image of the large text set as shown in Eq. (1),

$$f : S \rightarrow s||s| \ll |S| \tag{1}$$

where S represents a large amount of text collected and s represents the text document that is summarized. A large amount of text collected S size is greater than the size of the document summarized. The algorithm which performs the text summarization task is known as a text summarizer. Figure 1 depicts text summarization. Text summarizer is mainly divided in two categories, namely multi-document and single-document summarizers. In summarization of multi-document, a collection of text documents (i.e., multi-documents) is converted into a summary of a single document which indicates the total glimpse of these several documents, whereas in the summarization of single-document, a single huge document of text is converted to another summary of a single document.

The summarization of multi-document is an approach utilized to summarize several text documents and also used to understand a huge collection of text documents. Throughout multi-document summarizing, relevant sentences are extracted from a document collection based on the document content, resulting in a concise summary. In the current years, researchers are giving more attention to evolving techniques of document summarization. Numerous summarization approaches are proposed to create a summary from the given set of documents by removing the significant sentences. The summarization of multi-document is applied for understanding and examining a large set of documents, and the major resource of these sets are blogs, news archives, tweets, research papers, web pages, technical reports, and web search results available on the web and other areas. Few examples of the multi-document summarization applications are examining the results of web search for supporting users to browse further, and producing a summary for news articles.

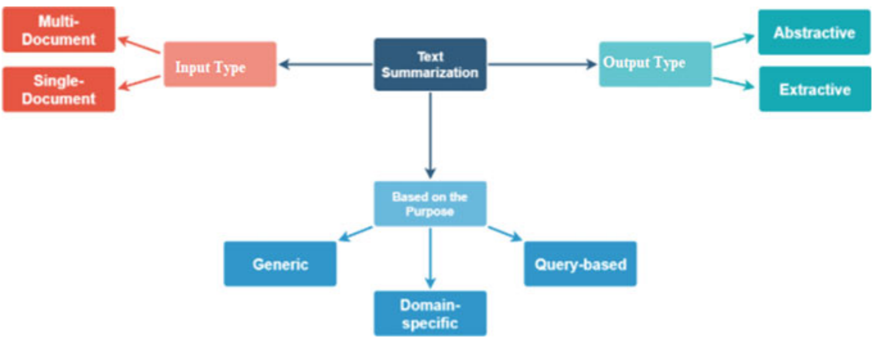


Fig. 1 Text summarization

In a huge collection of text documents, the generation of summary and processing of the document is a computationally difficult task. Especially, in the analytics of the big data era, where a collection of data size is tremendous thus the algorithms are needed for rapidly summarizing the large collections of text.

Summarization of a single document is easier to handle because it needs one text document to be examined for summarization, whereas handling summarization of multi-document is a difficult and complex task. It needs numerous (multiple) text documents to be examined for generating a meaningful (informative) and compact summary. When the number of documents in multi-document summarizing increases, the summarizer will find it increasingly harder to complete the task. A summarizer is preferable because it provides a more relevant and productive condensed representation of large amounts of content. Considering terms of semantic similarity will give advantages by creating a relevant summary. However, it is computationally costly, as the semantic concepts are analyzed and created from a huge collection of text in order to produce a summary. Thus, the issues with text summarization of single-document and multi-document are presented with the aid of modern techniques in text analytics.

2 Literature Review

Today, massive information is available on the web. It is difficult to gather significant data from this tremendous volume of information. In the case of text documents, it is an exhaustive and complex process to perceive and collect the prime data from a large volume of resources at an appropriate time for humans. Fortunately, the process is performed automatically permissible by data retrieval techniques for decades. However, increasing the quantity of data creates some performance problems like inadequate solutions, clumsy applications of data retrieval works, and so on. Machines with advanced technologies are used to decrease the destruction created by these problems. Thus, text summarization is applied to create a short form of documents by conserving the basic information. Several techniques and tools have been utilized for text summarization. Table 1 describes the details about the techniques handled by text summarization.

3 Conclusion

The growing popularity of the Internet has resulted in the availability of a significant amount of data. It is complex for people to summarize huge volumes of text. Thus, it needs automatic summarization equipment in this period of data overloading. Several extractive techniques for the summarization of single and multiple documents are emphasized. Many possible instances of summarizing the many documents are comparatively studied. The performance of summarization is effective, while both

Table 1 Text summarization

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
1	Mutlu et al. [6]	The dataset obtained will contain publications of academic and abstracts	The ensemble feature space is proposed on a model of a remote short-term memory-based neural network	In the present dataset, a comprehensive feature space is generated and an investigation is made for the effect of features on the resulting summary	For extractive and abstractive summarization tasks, a novel dataset is presented. The extract obtained is calculated to make sure the validity process of the human extraction of texts. Then, re-evaluation on the problem of extractive summarization is carried out on the present dataset of summarization. Moreover, to generate an informative summary, the feature vector is analyzed
2	Widyassari et al. [7]	85 conference and journal publications	A systematic and broad analysis of research in the text summarization is made, which is published in the year 2008–2019	It provides an explanation of the trends/topics in the text summarization field. And, also give references to features, preprocessing, and public datasets that are used	The approaches that are viewed from the results of the summary are abstractive and then extractive. The summary of extractive is heading to maturity and therefore, the study is shifted to real-time summarization and abstractive summation
3	Hou et al. [8]	Chinese dataset	An unsupervised technique that contains a deep neural model, domain knowledge, and rhetorical theory for automatic text summarization (ATS)	The results of the summary are evaluated without a golden standard, therefore unsupervised metrics of evaluation are proposed, whose hyper-parameters are regulated by supervised learning	It contains 3 components namely: (i) Domain knowledge is utilized for unsupervised parsing of rhetorical. Thus for every document rhetorical tree structure is derived. (ii) In the model of rhetorical parsing, the concept of translation is perceived to reduce the data scarcity issue. (iii) The subroutine model is built on the rhetorical tree's derived structure and it will produce content-balanced outcomes on summarization of text
4	Liang et al. [9]	LCSTS (i.e., social media dataset)	For abstractive text summarization of social media, a reinforced Seq2Seq (i.e., sequence-to-sequence) method is applied	To directly optimize the ROUGE score, reinforcement learning and cross-entropy are combined. The proposed method is **2.6%**, **2.5%**, and **2.1%** higher than the primary Seq2Seq model on the ROUGE-1, 2, and L, F1 score, respectively	The objective of abstractive summarization is to produce a concise version of a sentence while still conveying the core content. For filtering invalid information, a selective gate is added behind the encoder module. Though some approaches are attained remarkable results depending on the framework of sequence-to-sequence, there are still several issues that cannot be avoided

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
5	Cao et al. [10]	Nil	2-layer semantic link network (SLN) model	It shows that a summary composed by *paragraph* or *group* tends to consist of more phrases or keywords than a summary composed of sentences. Then, the summary contains *groups* which consist of more phrases or keywords than *paragraphs* especially while the source text length is from 7000 to 17,000 words, which is the actual scientific paper's length	A *framework of* text summarization *based on group* by using the SLN model is proposed and it will rank the groups and the unique text is concatenated into a summary. The semantic links used are *sequential* link, *is-part-of* link, *cause-effect*, and *similar-to* link
6	El-Kassas et al. [11]	DUC-2001 and DUC-2002 datasets	EdgeSumm	EdgeSumm gives a good ROUGE result on dataset DUC-2001. And for DUC-2002, the outcomes display that the present system outperforms the modern ATS models by attaining a good score of 4.7% and 1.2% for the ROUGE-L andROUGE-1metrics, respectively	A novel "EdgeSumm" framework integrates an extractive ATS set models (namely statistical, graph, centrality, and semantic-based models) to solve their sole drawbacks. EdgeSumm is not restricted to a particular field and it is unsupervised therefore it does not need training data

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
7	Mohd et al. [12]	DUC-2007 dataset	Distributional semantic model	When semantics is incorporated as a feature in text summarization, the results will be better, and it will help to reduce redundancy even more	Though existing automatic summarizer produces high-quality summary, they will not preserve the underlying semantics and meaning of the text The technique will acquire and preserve the text semantics as the primary attribute in this document summarization
8	Elbarougy et al. [13]	Nil	Al-Khalil morphological analyzer and modified page rank technique	The performance of the proposed approach is calculated by EASC Corpus. It will perform effectively with the 10,000 iterations	The Arabic language contains a complicated morphological structure and it is more difficult for extracting nouns that can be used as an attribute for the process of summarization. The present approach is depending on a graph model, which denotes the documents as a graph with the vertices representing sentences. The modified page rank technique uses numerous iterations to identify the better summary results, and the summary which is extracted mainly depends on the ratio of compression
9	Moradi et al. [14]	Nil	Bi-directional encoder representations from transformers with a clustering method	The method yields advanced results without demanding labor-intensive knowledge bases or annotated features creation or computationally requiring pre training of domain-specific. As a result, the investigation will serve as a basis for further research on biological text summarizing	In biomedical summarization of text, the main challenge is capturing the text context. The method will show how the deep bidirectional language approach produces contextualize embeddings and also in text summarization of biomedical how it is used to quantify the informative sentences

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
10	Yang et al. [15]	Two benchmark datasets	Plausibility-promoting generative adversarial network for abstractive text summarization with multi-task constraint	The tasks will aid to increase the quality of locating information and produce prime summaries using language modeling by reducing the problems of duplicating words and insufficient sentences. It attains reliable performance than the advanced baseline approaches in terms of the two qualitative and quantitative evaluations	The PGAN-ATSMT method will train a *G* generative and *D* discriminative model by adversarial learning. The G generative method will take the original text as input and generates a summary. A discriminator *model* will differentiate the summary produced by *G* from the ground truth summary by avoiding saturation. The generative and the discriminative approach will tune *G* to give plausible and finest summaries. Then, the generative model is proposed by utilizing the multi-task learning, and also shares their LSTM decoder and encoder with syntax annotation and text categorization task, respectively
11	Zaman et al. [16]	EurekaAlert	(Hybrid architecture of abstractive and extractive summarization (HTSS)	The HTSS model will outperform abstractive summarization of text on the score of ROUGE and neural simplification of text on the SARI score	Whereas simplification will reduce the document complexity, summarization will attempt to decrease the document length, by retaining the original meaning. The author extended the generator approach for the combined work of simplification and summarization
12	Uçkan et al. [17]	Document Understanding Conference (DUC-2004 and DUC-2002) datasets	This method proposed is a novel approach for generic, and extractive text summarization	In the summary, it will prevent word group repetition. The developed method reached a 0.59208 ROUGE value for a 400-word summary, a 0.5195 value for 200 words, and 0.38072 for 100-word summaries	The maximum independent set is used in text summarization. In addition, the KUSH tool for processing a text is suggested to retain the semantic cohesion among sentences. The anticipation is that the group of sentences should be removed from the summary, which is representing to the nodes in the set of independent. Thus, the nodes which are forming the set of independent on the graph are detected and excluded from the graph

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
13	Rouane et al. [18]	Randomly selected 100 papers of biomedical from the *Bio Med Central database full-text*	Data mining approaches: frequent item set mining and clustering	The obtained summary is compared with 5 known summarizers: *Text Teaser, Text Rank, SweSum, Microsoft AutoSummarize*, *Item Set*, and also with two baselines, i.e., summarization by applying only the frequent itemsets mining (FRQ-CL), and summarization by applying only the CL-FRQ (clustering)	The biomedical paper contains biomedical ideas set by applying the UMLS meta thesaurus. Similar sentences are clustered by applying the K-means algorithm. The A priori technique is used to find the frequent item sets between the clustered sentences. The salient sentences from each cluster are then selected using the detected frequent item sets to generate the summary
14	Sanchez-Gomez et al. [19]	Document understanding conference (DUC)	Multi-objective artificial bee colony algorithm (MOABC)	The asynchronous design will improve the parallel design largely, by being greater than 55 times speedy with 64 threads than the common model. The 86.72% efficiency is noted for 64 threads	Because of the continual development of information on the web, the automatic multi-document summarization of text is the main task in various knowledge fields. Following a time analysis of the technique, a runtime comparison is made between the use of several random number generators in the design. A MOABC algorithm is designed by following its native scheme, in which parallelization is implemented to the main steps, and then several schedules of parallel are compared and studied. Finally, another design that depends on the asynchronous nature of the bee colony is compared and implemented

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
15	Joshi et al. [20]	Nil	SummCoder	To assist law enforcement agencies, the TIDSumm (Tor Illegal Documents Summarization) dataset is introduced. It comprises two ground truth summary sets, which are manually constructed, for hundreds of web documents retrieved from onion websites available in the Tor system	A new method for generic extractive summarization of text in single documents. The method will provide a summary based on 3 selection metrics of sentence formulated by: sentence novelty, sentence content, and position relevance. The content metric of a sentence is calculated by applying a deep auto-encoder system, and the metric of novelty is determined by employing the similarity between sentences which is indicated as embeddings in a semantic space. Then, the position metric is a type of hand-designed attribute, which will assign additional weight to initial sentences by calculating dynamic weight operation that is managed by the document length
16	Mutlu et al. [21]	DUC dataset	Fuzzy inference method	The performances of summarization are calculated by utilizing performance metrics of actual classification. The model recommended the fuzzy system's usage depending on a fuzzy rule and a feature vector for extracting text summarization. By changing the ratio of compression, the extraction models are calculated	The study mainly focused on the extraction of informative summaries from many documents by utilizing handcrafted attributes from the literature. Initially, an investigation is made on the feature vector generation. Some of the features are term frequency, several sentences, title similarity, sentence position, sentence to sentence similarity, sentence length, sentence phrases, bushy-path outcomes, proper nouns, documents length, and n-gram co-occurrence. Then, an examination of several features combinations is carried out. The fuzzy inference model and a shallow multi-layer perceptron are applied to extract essential sentences from texts in the dataset of DUC

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
17	Hark et al. [22]	DUC-2004 and DUC-2002 datasets	Karcı summarization	The proposed summarizer will outperform all modern methods by a 200-word summary in the ROUGE-2, ROUGE- W-1.2, ROUGE-L, and ROUGE-1 metrics. Additionally, it outperforms the nearby competitive summarizers by a 6.4% factor for ROUGE-1 on the dataset of DUC-2002	A new method is utilized for extractive and generic text summarization. Initially, the karcı entropy is used in a summarization of document approach. The proposed model does not take any training data or information source. TheKUSH tool is introduced for processing input text and it also protects the consistency of semantic among sentences. Thus the karcı entropy solution will select the generic and informative sentences in a text unit or paragraph
18	Goularte et al. [23]	Brazilian Portuguese text dataset	Fuzzy rules	The proposed model for summarization of text with a few number of fuzzy rules will aid in the growth and usage of a future expert model automatically evaluate writing. The proposed approach is made a comparison with further techniques including a score, model, sentence, and naive baseline, by applying ROUGE measures. Thus the proposal gives good f-measure (95% CI) score than the aforementioned models	An automatic procedure is presented for text assessment which is based on fuzzy rules on different types of extracted attributes that identify the significant data in the text assessment. The comparison is made between text summary which is produced automatically and with reference summary that is generated by domain experts. Then, the method makes summaries on a text by examining correlated features to decrease dimensionality, and fuzzy rules are also applied for the summarization of text

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
19	Mohamed et al. [24]	DUC-2002	Semantic role labeling (SRL) and explicit semantic analysis (ESA)	The testing performed depicts the system scalability, i.e., by changing the evaluation size of data which shows the performance of the summarizer with small impact, especially for the task of summarization on a single document. Finally, in a summary, the findings display the power of the vectorial semantic and role-based representation while it is integrated with the crowd-sourced knowledge in Wikipedia	A creative text summarizing structure based on a graph is proposed for the generic sole and multiple documents. The summarizer is benefited from text semantic representation models: SRL and ESA and also from the constantly growing collective knowledge of humans in Wikipedia. Using the ESA paradigm, the SRL is applied to parse sentence semantics, where a word's tokens are denoted as a vector of weighted Wikipedia ideas. The essence of the proposed structure is to build a graph representation concept that is supported based on the semantic role for summarization
20	Mosa et al. [25]		Swarm intelligence (SI) optimization technique	The swarm intelligence (SI) has never been utilized to enhance real-time summary models previously. Thus, a novel approach is proposed to be ample for achieving multiple objectives and also to fulfill the need of real time. And, the method will help researchers to consider the several kinds of SI while resolving the summarization tasks, specifically, in the short text summarization (STS) field	To resolve the text summarization work effectively, the optimization model called swarm intelligence is used. Unfortunately, 3 kinds of text summarization work by utilizing SI indicate bit in the study while contrasted with the further summarization models such as genetic algorithm and machine learning. It has been noted that the task of summarization with various types is not been regulated as a multi-objective optimization (MOO) task before, though multiple objectives are considered

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
21	Tsai et al. [26]	Nil	Sentiment-based techniques	The systematic model initially built classifiers to find useful reviews and the sentences are classified into 6 hotel features within the helpful reviews. At last, it analyze the sentiment polarity of sentences to create the summaries of the review	In today's hotels, the manual process of review readers is difficult in the level of online reviews. Automatic summarization of review is a promising direction to improve the detail processing of travelers. To compile a summary of the review, the author focused to perform sentiment analysis or extract relevant features of a text. However, many reviews consist of non-sentimental or non-specific content, by hindering the sentiment techniques' in which it will precisely summarize useful detail from hotel reviews

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
22	Nawaz et al. [27]	Nil	Local weights (LW) and global weights (GW) approach	The summary extractive is created by LW- and GW-based models. It is calculated with a ground truth summary which is acquired by the specialists. Thus, the VSM is applied as a standard structure for sentence weighting. Investigation display that the LW model is desirable for the generation of extractive summaries. The F-score value of the sentence weight and the weighted term-frequency approach are 80% and 76%, respectively. The VSM has reached an accuracy of 62% on a similar dataset. The code and ground truth datasets are made publicly accessible to researchers	Currently, many advancements are made in Urdu linguistics. There is no publicly accessible dataset or structure for automatic Urdu extraction summary generation. The sentences which have the strongest weights are given more significance to be noted in the summary. Then, the weight of a sentence is calculated by the total of the weights of the words in the sentence. The two well-known models are applied to calculate the weights of the words n the English language are local weights (LW) and global weights (GW) models. The weight sensitivity based on the text contents, the single word will have various weights in several articles, it is called as LW model. Whereas, in the GW case, the word weights are calculated from the free dataset, which indicates the weights of every word will stand similar in several articles. Thus in the present framework, GW and LW methods are prepared for the language of Urdu. The weighted term-frequency and sentence weight approach are LW-based models. It will calculate the sentence weights by the total of salient words and the total of frequencies of the salient words, respectively. The vector space model (VSM) is based on the GW model. GW is largely utilized in the English language for several applications like text classification and information retrieval

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
23	Sanchez-Gomez et al. [28]	DUC datasets	Multi-objective artificial bee colony algorithm based on decomposition (MOABC/D) algorithm	The MOABC/D is applied to resolve the extractive text summarization of multi-document problems. An asynchronous design of MOABC/D method is applied to take benefit of multi-core frameworks. The results obtained are enhanced the prevailing ones in the scientific report for ROUGE-2, ROUGE-1, and ROUGE-L scores as well report a better speedup	Because of the overflow of textual detail on the web, automatic summarization of text techniques is becoming significant in various knowledge fields. Extractive text summarization of multi-document models will automatically create summaries from a collection of documents, also cover the important content, and avoids redundancy of information. These models are addressed by using optimization models. In the scientific report, many are single-objective optimization models, but currently, multi-objective models are evolved and also enhanced the prevailing results of single-objective. Additionally, in the multi-objective optimization field, the approaches based on decomposition are applied successfully
24	Anand et al. [29]	Nil	Neural network	The major benefit of the present models is that they do not depend on domain-specific knowledge, or handcrafted attributes, nor their application is limited to a specific sub-domain thus preparing them preferable to be expanded to further fields as well. The unavailability issue of labeled data is tackled for the work by allocating scores/classes to sentences in the set of training, depending on their parallel with a summary of reference which is generated by humans	The analysis documents which are available in digital form will offer various opportunities for the extraction of information and application. Because of the more complexity and unusual shape of the documents, the automatic summarization of the legal texts is a more pivotal and difficult task. Previous models rely on large labeled datasets, utilizing hand-engineered attributes, holding domain knowledge and their observation is focused on a narrow sub-domain to increase effectiveness. The simple generic models are proposed applying neural networks for the task of summarization for Indian valid judgment documents. The two neural network designs are explored for this work using the sentence and word embeddings for catching the semantics

(continued)

Table 1 (continued)

S.No.	Reference	Dataset size	Techniques	Conclusion	Remarks
25	Leiva et al. [30]	Nil	Text rank algorithm	The RTS feasibility is validated on blog sites. It demonstrates that the performance of runtime incurred in the present method is negligible. The method shows that time for reading is decreased by a factor of four without affecting largely the quality of the summary. Generally, anybody interested in creating their web content will concise but informative can benefit from this task	Responsive text summarization (RTS) is an approach to web design. Its goal is to permit users to read desktop web pages based on the dimension of the device they are using. RTS applies the Text rank technique, so it is exploited to produce short summaries (apt to smart phones, where it has premium screen space) or longer, also explicative summaries (more apt, e.g., for laptops or tablets)

semantic similarity and clustering are considered. Regarding semantic similarity, it provides better pyramid scores, ROUGE, and retention rates for a summary. In the future work, this will provide the support to multi-lingual summarization of text to facilitate the generation of summary from the collections of text documents that are available in various languages.

References

1. Alami N, Meknassi M, En-nahnahi N (2019) Enhancing unsupervised neural networks based text summarization with word embedding and ensemble learning. Expert Syst Appl 123:195–211
2. Sanchez-Gomez JM, Vega-Rodriguez MA, Perez CJ (2020) Experimental analysis of multiple criteria for extractive multi-document text summarization. Expert Syst Appl 140:112904
3. Wang WM, Li Z, Wang JW, Zheng ZH (2017) How far we can go with extractive text summarization? Heuristic methods to obtain near upper bounds. Expert Syst Appl 90:439–463
4. Moradi M (2018) CIBS: A biomedical text summarizer using topic-based sentence clustering. J Biomed Inform 88:53–61
5. Moirangthem DS, Lee M (2020) Abstractive summarization of long texts by representing multiple compositionalities with temporal hierarchical pointer generator network. Neural Netw 124:1–11
6. Mutlu B, Sezer EA, Akcayol MA (2020) Candidate sentence selection for extractive text summarization. Inf Process Manag 57(6):102359
7. Widyassari AP, Rustad S, Shidik GF, Noersasongko E, Syukur A, Affandy A (2020) Review of automatic text summarization techniques & methods. J King Saud Univ Comput Inf Sci
8. Hou S, Lu R (2020) Knowledge-guided unsupervised rhetorical parsing for text summarization. Inf Syst 94:101615
9. Liang Z, Du J, Li C (2020) Abstractive social media text summarization using selective reinforced Seq2Seq attention model. Neurocomputing 410:432–440
10. Cao M, Zhuge H (2020) Grouping sentences as better language unit for extractive text summarization. Fut Gen Comput Syst 109:331–359
11. El-Kassas WS, Salama CR, Rafea AA, Mohamed HK (2020) EdgeSumm: graph-based framework for automatic text summarization. Inf Process Manag 57(6):102264
12. Mohd M, Jan R, Shah M (2020) Text document summarization using word embedding. Expert Syst Appl 143:112958
13. Elbarougy R, Behery G, El Khatib A (2020) Extractive Arabic text summarization using modified PageRank algorithm. Egypt Inform J 21(2):73–81
14. Moradi M, Dorffner G, Samwald M (2020) Deep contextualized embeddings for quantifying the informative content in biomedical text summarization. Comput Methods Programs Biomed 184:105117
15. Yang M, Wang X, Lu Y, Lv J, Shen Y, Li C (2020) Plausibility-promoting generative adversarial network for abstractive text summarization with multi-task constraint. Inf Sci 521:46–61
16. Zaman F, Shardlow M, Hassan SU, Aljohani NR, Nawaz R (2020) HTSS: A novel hybrid text summarisation and simplification architecture. Inf Process Manag 57(6):102351
17. Uçkan T, Karcı A (2020) Extractive multi-document text summarization based on graph independent sets. Egypt Inform J 21(3):145–157
18. Rouane O, Belhadef H, Bouakkaz M (2019) Combine clustering and frequent itemsets mining to enhance biomedical text summarization. Expert Syst Appl 135:362–373
19. Sanchez-Gomez JM, Vega-Rodríguez MA, Perez CJ (2019) Parallelizing a multi-objective optimization approach for extractive multi-document text summarization. J Parallel Distrib Comput 134:166–179

20. Joshi A, Fidalgo E, Alegre E, Fernández-Robles L (2019) SummCoder: An unsupervised framework for extractive text summarization based on deep auto-encoders. Expert Syst Appl 129:200–215
21. Mutlu B, Sezer EA, Akcayol MA (2019) Multi-document extractive text summarization: a comparative assessment on features. Knowl Based Syst 183:104848
22. Hark C, Karcı A (2020) Karcı summarization: a simple and effective approach for automatic text summarization using Karcı entropy. Inf Process Manag 57(3):102187
23. Goularte FB, Nassar SM, Fileto R, Saggion H (2019) A text summarization method based on fuzzy rules and applicable to automated assessment. Expert Syst Appl 115:264–275
24. Mohamed M, Oussalah M (2019) SRL-ESA-TextSum: a text summarization approach based on semantic role labeling and explicit semantic analysis. Inf Process Manag 56(4):1356–1372
25. Mosa MA, Anwar AS, Hamouda A (2019) A survey of multiple types of text summarization with their satellite contents based on swarm intelligence optimization algorithms. Knowl Based Syst 163:518–532
26. Tsai CF, Chen K, Hu YH, Chen WK (2020) Improving text summarization of online hotel reviews with review helpfulness and sentiment. Tour Manag 80:104122
27. Nawaz A, Bakhtyar M, Baber J, Ullah I, Noor W, Basit A (2020) Extractive text summarization models for Urdu language. Inf Process Manag 57(6):102383
28. Sanchez-Gomez JM, Vega-Rodríguez MA, Perez CJ (2020) A decomposition-based multi-objective optimization approach for extractive multi-document text summarization. Appl Soft Comput 91:106231
29. Anand D, Wagh R (2019) Effective deep learning approaches for summarization of legal texts. J King Saud Univ Comput Inf Sci
30. Leiva LA (2018) Responsive text summarization. Inf Process Lett 130:52–57

IoT-Enabled Applications for Elderly Support and Care: A Systematic Review

Shrija Madhu, K. Jyothi, S. Pravallika, and N. Sridurga

Abstract Around the world a major challenge today is taking care of the aging population. Old age homes are expensive, and many would prefer to stay in their own house. The main issue here is that the younger generation may have to work to meet the family needs, leaving the elderly at home. It raises a lot more issues like safety, care, and support to the elderly. In some cases though home nurses are available, but the family may not be in a position to afford the extra expense of hiring a home nurse. Internet of Things (IoT) is a boon for such elderly people. This paper focuses on all such aspects and brings forth a plethora of options through IoT devices, for the young generation to take care of the elderly. We have divided the IoT devices for the elderly into three major areas like safety and comfort, monitoring, and treatment related devices. The analysis done in this paper could be useful for developing better and useful IoT-based healthcare devices for support and care of the elderly people.

Keywords IoT · Healthcare · Monitoring devices · Safety devices · Devices for comfort

1 Introduction

The Internet of Things (IoT) is a collection of digital devices that are interrelated and wireless. These devices can collect, send, and store data over a network without human intervention. Around the world, policy makers are implementing policies to deliver healthcare services using IoT-based and related technologies. IoT devices are making rapid entry into the field of healthcare, providing all round services like predicting health issues, diagnosing, treatment, and monitoring patients [1–5]. The benefits of IoT have made researchers and developers dig deep in search how the

S. Madhu (✉) · K. Jyothi · S. Pravallika · N. Sridurga
Godavari Institute of Engineering and Technology(A), Rajahmundry, India
e-mail: shrija@giet.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_71

emerging IoT technologies can support health systems to deliver safe and effective care.

IoT for healthcare can be just a device that can connect to the Internet and can collect health-related data. It can collect and measure data from individuals, computing devices, smart or mobile phones, or other portable devices [2, 3].

Recent studies on the services and applications of IoT in healthcare have shown its usage in the form of e-Health, m-Health, wearable and semantic devices, and smart gadgets [4]. IoT provides the ability to track and monitor health progress remotely by healthcare professionals. It also helps in fast-track symptom identification, clinical diagnoses, and remote treatment in case of emergencies. IoT devices also assist in the early detection of abnormalities and early intervention in chronic cases [5]. Moreover the IoT-based applications provide better quality and low-cost medical care.

IoT is one of the smart technologies to integrate smart devices on networks such as the big data, cloud computing, and artificial intelligence to make the healthcare more efficient, convenient, and personalized [6]. IoT is going to be future network for various application fields which include real-world scenario of smart autonomous hospital management comprising the technical operations like remote device coordination, the architecture of network, middleware, databases, and other application services [7].

2 Background and Motivation

Healthcare is the most expensive in the world. Health apps have the potential to improve efficiency and value while lowering costs. Currently, 3,25,000 health apps have been developed. If the health apps are to be successful and broadly adopted, patients and clinicians must have confidence that they are safe and effective [8]. The benefits of health applications using IoT are remote health monitoring, fitness programs, child care, and elderly care [9, 10].

IoT-based healthcare and use of deep machine learning for the analysis and diagnosis of biomedical and healthcare problems are the recent trends. The deep learning (DL) algorithms had high sensitivity and specifying capacity for detecting diabetic retinopathy and macular edema in retina. The DL had cascaded into the network and made use for medical big data. Present biological and medical devices are implemented by using the concept of big data. DL can also provide the analysis for healthcare apps [11, 12]. The deep learning includes deep care for disease progression modeling and future risk prediction. Convolutional neural network (CNN) is one of the classic deep learning models and deep neural network based approach for patient data de-identification. Some examples of this include osteoporosis prediction, the brain dynamics by electroencephalography (EEG), detecting and classifying cardiac arrhythmias [13–18].

In 2020 to mitigate the spread of COVID-19, the governments, policy makers, and healthcare professionals used virtual models and some of the smart devices like mobile sensor data, cross-validated by other big sensor surveillance data, geopositioning method to identify the potential contact persons [19].

IoT is also changing the way enterprises are offering products and services to end users. The volume of IoT devices is going to increase exponentially in every country, which will be the future driving force in the development of science and technology. Majority of hospitals use IoT services for asset management. Some medical devices like sensor diagnostic and imaging devices are in use and are considered as smart devices [20]. Australia is also in the process of establishing a policy for IoT development and investment [21]. There are self-healing mechanisms which also can be employed using IOT-based techniques [22].

3 IoT Architecture for Healthcare Applications

The framework of IoT is applied to healthcare applications to connect different medical devices through various communication channels. The architecture followed for healthcare and especially the devices for elderly in Fig. 1 has three layers: sensing layer, fog layer, and the cloud layer. Medical devices such as pulse-oximeter, fluid sensor, electrocardiogram, sphygmomanometer, thermometer and so on are called sensing elements. These are directly or indirectly connected to the patients and will monitor and sense various physical parameters related to the patient's health. The information sensed by the sensing elements are transmitted to the next layer called processing layer. Based on the distance between the sensing layer and the processing layer, different communication protocols are used such as Zigbee, Wi-Fi, Z-wave, and MQTT.

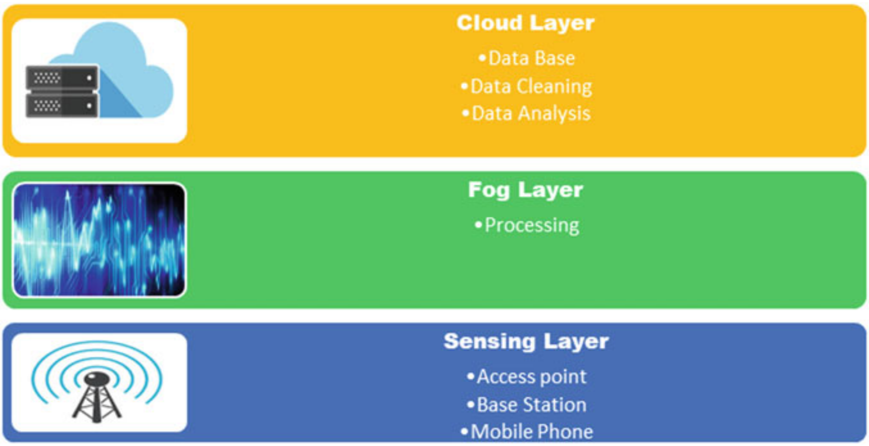


Fig. 1 IOT architecture for the elderly healthcare

As there is consistent release of data from sensing devices, which may be critical or time sensitive in such cases, it is necessary to take immediate action and quick response to save the life of the patient. To address such problems, data collected by the sensing elements is processed by the devices at the fog computing layer. The idea behind fog computing is to process and store data as close to sensing devices as possible. Later the same data is sent to the cloud where it will be stored for further analysis. As the magnitude of data sent by the sensing elements is huge, it will be stored in a cloud which will act as a repository. This repository acts as a source for different disease analysis [23].

4 IoT Devices for the Elderly

Elderly people often struggle to master new technology, but a recent study found that around 80% of senior citizens over the age of 65 own a regular phone while 20% use smartphones. Due to the COVID-19 pandemic, smartphone use has gone from a luxury to a necessity, forcing older adults to use telemedicine solutions to reduce risk. Moreover, 25% of seniors now use health apps on their smartphones, and 28% monitor their health with wearables. The statistics above show that modern senior citizens are tech-savvy enough to learn and use smart solutions [24, 25].

IoT devices for the elderly can be classified into three categories:

1. Devices for Monitoring
2. Devices for Treatment
3. Devices for Safety and Comfort.

Figure 2 shows the classification of these devices. Table 1 gives details of the commonly available monitoring, treatment, and devices for safety and comfort [22, 24–27].

4.1 Healthcare Monitoring

IoT devices present a number of novel opportunities for patients to self-monitor and for healthcare professionals to monitor patients. As a result, both healthcare providers and their patients can benefit from the range of wearable IoT devices.

Remote patient Monitoring. In healthcare, remote patient monitoring is the most significant one. In healthcare centers, these IoT devices are used to gather patient's health data automatically, even if the patient is not physically present in the health center.

Glucose Monitoring. IoT gadgets can assist in overcoming these difficulties by continuously and automatically monitoring patients' blood and glucose levels. In

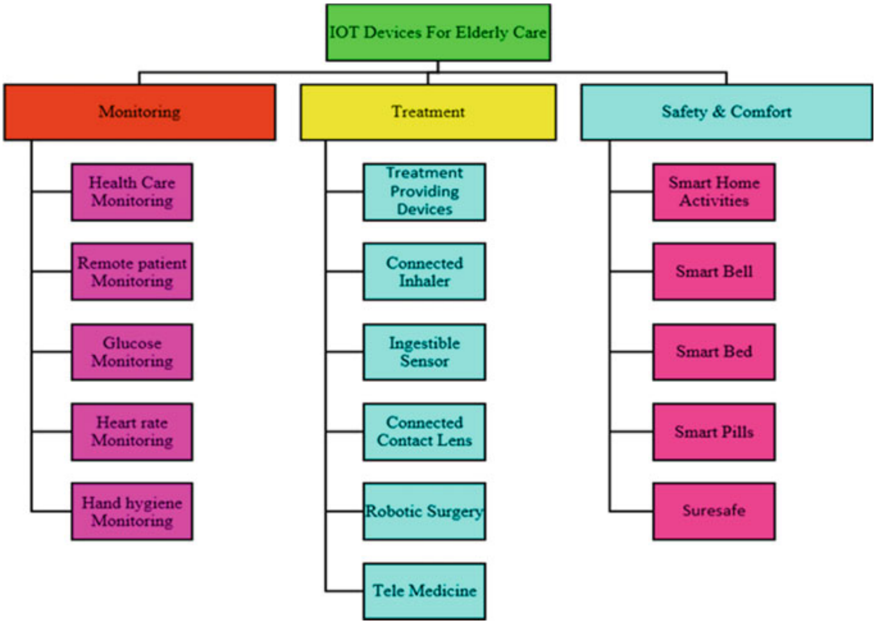


Fig. 2 Classification of IOT devices for elderly

Table 1 Types of devices and the available devices under each

S. No.	Type of device	Available devices
1	Monitoring devices	Glucose monitoring, heart-rate monitoring, hand hygiene monitoring,wearable health devices
2	Treatment devices	Connected inhaler, ingestible sensor, connected contact lens, and robotic surgery
3	Safety and comfort devices	Smart bell, smart bed, smart pills, and SureSafe

addition to removing the need for manual record keeping, glucose monitoring devices can warn patients when their blood sugar levels are troublesome.

Heart-rate Monitoring. Even for people who are present in healthcare facilities, measuring heart rates might be difficult, similar to glucose. Regular heart rate checks cannot protect against sudden changes in heart rate, and traditional continuous cardiac monitoring techniques used in hospitals entail patients being permanently hooked to machines, which limits their mobility.

Hand hygiene Monitoring. When patients enter hospital rooms, many hospitals and other healthcare facilities employ IoT devices to remind them to wash their hands. The gadgets can even provide guidance on the optimal methods of sanitization to reduce a specific risk for a certain patient [26, 27].

4.2 Treatment Providing Devices

There are IoT devices that go beyond monitoring and actually provide treatment, or even live in or on the patient, even if wearable devices like those mentioned above continue to be the most popular type of IoT device in healthcare.

Connected Inhaler. Attacks caused by diseases like asthma or COPD frequently occur unexpectedly and without much notice. IoT-connected inhalers can benefit patients by tracking the frequency of attacks and gathering environmental data to help medical professionals determine what precipitates an attack.

Ingestible Sensor. Ingestible sensors make it possible to gather data from the digestive and other systems considerably less invasively, e.g., they can help identify the cause of internal bleeding or offer information into stomach PH levels.

Connected Contact Lens. Another chance to gather healthcare data passively and unobtrusively is offered by smart contact lenses. The fact that connected contact lenses have been patented by businesses like Google suggests that they may also have tiny cameras that would enable users to shoot images with their eyes.

Robotic Surgery. Robotic devices are used for performing surgery without human intervention. These IoT devices are understanding the internal workings for performing the surgery effectively inside the bodies.

Telemedicine. The idea behind telemedicine is to provide people who live in remote locations, especially those without medical facilities, with remote healthcare services. Additionally, it served to lessen the strain on hospitals that are overrun with patients [24, 25].

4.3 Safety and Comfort

Smart Home Activities. This group of intelligent devices tracks a person's location and tracks movement patterns as well as falls. In case, there has not been any movement for a long time, caregivers can install a motion sensor that sends a warning.

Smart Bell. The security in this system combines the capabilities of a smartphone and a home network system. It allows customers to keep an eye on guests in real time from a distance using an IoT-based doorbell that is placed next to a home's front door. By automatically collecting the image, processing it for facial recognition, and communicating it to the server via email to determine if the intruder is known or unknown, this method increases security's autonomy. Additionally, you can make them known by inputting the OTP that is delivered to the mail server.

Smart Bed. With the latest trends in IoT there is a new variety of bed called smart bed with many adjustable features. A smart bed usually has features like mattress adjustment for position, temperature, and firmness.

Smart Pills. Despite the fact that many people desire to take their prescriptions correctly, obstacles frequently stand in the way. It may be challenging to read small labels and distinguish between different tablets when your vision deteriorates. Another problem is that it is simple to forget to take prescriptions or to skip a dose. For the elderly support, these smart pills help to take the medicine at the specified time by giving an alert and displaying the medicine to be taken. Users will get 30 min time to take the medicine.

SureSafe. A wearable device that automatically alerts a caregiver when a fall is detected [25].

5 Conclusion

The purpose of this research is to bring forth the IoT devices and applications useful in elderly healthcare. In the present study, three main categories of IoT applications/ devices in elderly healthcare are identified. The important finding of this research is that IOT for elderly is used 70–80% for monitoring, diagnosis, and treatment only. In healthcare, IoT applications are used nearly 70% for elderly healthcare. IoT usage for comfort in the form of smart homes and smart gadgets is mostly preferred by the younger generation. Moreover, IoT application in healthcare is around 65%, which shows that IoT along with other computing technologies like cloud computing, fog computing, and big data analytics are the powerful trending combinations.

Funding None.

Conflict of Interest There is no conflict of interest regarding publication of this paper from any of the authors.

References

1. Jaimon TK, Katrina LC, Enying G, Paul S (2020) The internet of things: impact and implications for health care delivery. J Med Internet Res 22(11):1–11
2. Internet of Things (IoT) in Healthcare, Research Markets (2019). https://www.medicaldevice-network.com/comment/bringing-internet-things-healthcare/. Accessed 02 Nov 2022
3. Dang LM, Piran MJ, Han D, Min K, Moon H (2019) A survey on internet of things and cloud computing for healthcare. Electronics 8(7):768

4. Nazir S, Ali Y, Ullah N, García-Magariño I (2019) Internet of things for healthcare using effects of mobile computing: a systematic literature review. Wirel Commun Mobile Comput 1–20:2019
5. Yin Y, Zeng Y, Chen X, Fan Y (2016) The internet of things in healthcare: an overview. J Ind Inf Integration 1:3–13
6. Tian S, Yang W, Grange JM, Wang P, Huang W, Ye Z (2019) Smart healthcare: making medical care more intelligent. Global Health J 3(3):62–65
7. Thangaraj M, Ponmalar P, Anuradha S (2015) Internet of Things (IoT) enabled smart autonomous hospital management system—a real world health care use case with the technology drivers. In: IEEE international conference on computational intelligence and computing research. IEEE, pp 10–12
8. Bates DW, Landman A, Levine DM (2018) Health apps and health policy: what is needed? J Am Med Assoc 320(19):1975–1976
9. Yu L, Lu Y, Zhu X (2012) Smart Hospital based on the internet of things. J Netw 7(10):1654–1661
10. Wisniewski H, Liu G, Henson P, Vaidyam A, Hajratalli NK, Onnela J, Torous J (2019) Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. Evid Based Ment Health 22(1):4–9
11. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. J Am Med Assoc 316(22):2402–2410
12. Tobore I, Li J, Yuhang L, Al-Handarish Y, Kandwal A, Nie Z, Wang L (2019) Deep learning intervention for health care challenges: some biomedical domain considerations. JMIR Mhealth Uhealth 7(8):e11966
13. Pham T, Tran T, Phung D, Venkatesh S (2017) Predicting healthcare trajectories from medical records: a deep learning approach. J Biomed Inform 69:218–229
14. Cheng Y, Wang F, Zhang P, Hu J (2016) Risk prediction with electronic health records: a deep learning approach. In: Proceedings of the 2016 SIAM international conference on data mining; SIAM'16. SIAM, Florida, USA, pp 432–440
15. Li H, Li X, Ramanathan M, Zhang A (2014) Identifying informative risk factors and predicting bone disease progression via deep belief networks. Methods 69(3):257–265
16. Lu N, Li T, Ren X, Miao H (2017) A deep learning scheme for motor imagery classification based on restricted boltzmann machines. IEEE Trans Neural Syst Rehabil Eng 25(6):566–576
17. Mathews SM, Kambhamettu C, Barner KE (2018) A novel application of deep learning for single-lead ECG classification. Comput Biol Med 99:53–62
18. Chen C, Jyan H, Chien S, Jen H, Hsu C, Lee P, Lee C, Yang Y, Chen M, Chen L, Chen H, Chan C (2020) Containing covid-19 among 627,386 persons in contact with the diamond princess cruise ship passengers who disembarked in Taiwan: big data analytics. J Med Internet Res 22(5)
19. Dauwed M, Meri A (2019) In: Ismail Y (ed) IOT service utilisation in healthcare. IntechOpen, New York, USA
20. Horizon Scanning Series: An overview. Australian Council of Learned Academies. https://acola.org/programs/horizon-scanning-series/. Accessed 19 Nov 2022
21. Mohammad MK, Turki MA, Amani AA, Faris AA (2022) IoT-based health monitoring system development and analysis. Secur Commun Netw
22. Shrija M, Harini NS, Vandana P, Jose C (2021) Self-healing of the body: the three basic key factors of self healing. Int J Innov Eng Manag Res 10(1):222–226
23. IoT applications in healthcare, wireless watchdogs a dataprise company. https://www.wirelesswatchdogs.com/blog/IoT-applications-in-healthcare. Accessed 16 Nov 2022
24. 7 of the Best IoT devices for senior citizens, help cloud (2018). https://www.helpcloud.com/blog/7-IoT-devices-for-senior-citizens/. Accessed 16 Nov 2022
25. IoT for seniors: how technology improves quality of life for the elderly. https://www.softeq.com/blog/IoT-for-seniors-how-technology-improves-quality-of-life-for-the-elderly. Accessed 16 Nov 2022

26. The ultimate guide to remote patient monitoring devices (2022). https://welkinhealth.com/remote-patient-monitoring-devices/. Accessed 16 Nov 2022
27. 10 Internet Of Things (IoT) healthcare examples, and why their security matters. https://ordr.net/article/IoT-healthcare-examples/. Accessed 16 Nov 2022

A Review on Federated Learning-Based Network Intrusion Detection System

K. Radhika, Bindhu Sree Reddy Alla, Sai Bhargavi Mamidi, and Nishanth Puppala

Abstract The security of network-connected devices is a subject of increasing concern due to the widespread use of the Internet in recent years and the possibility of numerous flaws that may be exploited by an attacker. Mobile phones, wearable technology, and self-driving cars are just a few examples of distributed networks that produce and transmit enormous amounts of data daily. The security and privacy of such devices are significantly enhanced by intrusion detection systems. For well-known intrusion detection systems (IDS) to detect increasingly sophisticated cybersecurity assaults effectively and efficiently, machine learning (ML) techniques must be applied. Due to their success in achieving high classification accuracy, these techniques have shown themselves to be quite useful. Thus, many solutions were developed to provide protection against cyberattacks and intruders. Many papers from many reputed authors were studied to understand the working of these solutions. However, the requirement to store and transmit data to a centralized server may put privacy and security concerns in jeopardy. Federated learning (FL), a privacy-preserving decentralized learning strategy that trains models locally and sends the parameters to a centralized server, fits in well to reduce privacy concerns associated with centralized systems. A computational methodology for networked machine learning called federated learning enables numerous cooperating organizations to train a single large-scale model. The rest of this review paper goes in-depth about the various solutions proposed by multiple authors. Their methodology, results, advantages, and disadvantages are analyzed, compared, and contrasted.

Keywords Network intrusion detection System · Federated learning

K. Radhika · B. S. R. Alla · S. B. Mamidi (✉) · N. Puppala
IT Department, Chaitanya Bharathi Institute of Technology, Hyderabad, India
e-mail: saibhargavi260@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_72

1 Introduction

Internet security is becoming a problem for businesses in the modern world. Web firewalls, encryption, authentication, and virtual private networks (VPN) have long been used to secure network infrastructure and Internet communication in the process of protecting data. The most recent addition to this group of security technologies is the intrusion detection system. The creation of IDS improves network security and safeguards organizational data. Any harmful behavior on the network is made visible to the network administrator with the help of the IDS, which then alerts the administrator, so they may take the necessary precautions to secure the data.

Our approach to an intrusion detection system is based on federated learning, where the server aggregates local model weights rather than the source data to protect user privacy. The result of using federated learning in this situation is to train a global model that achieves high detection accuracy across all edge devices involved.

The server acting as the central entity chooses the statistical model that will be taught. The trained model is then distributed to numerous decentralized computers or servers. These dispersed nodes use their own data to locally train the model. Finally, without gaining access to the data from the other nodes, the central entity pools the output from the spread nodes and creates a shared model.

2 Literature Survey

The goal of the study [1] is to pinpoint crucial decreased input properties for developing a successful and efficient IDS. Correlation-based feature selection, information gain, and gain ratio are used to assess the effectiveness of three common feature selection strategies. To discover significantly decreased input features, a feature vitality-based reduction method is proposed in this study. On smaller datasets, one of the effective classifiers, naive Bayes, is used to detect intrusions. The authors in [2] proposed an ensemble IDS for IoT devices. The suggested IDS employs the multi-agent system (MAS) concept and combines the J48 classification method with the information gain feature selection technique. The IDS's effectiveness is influenced by the categorization algorithm that is employed. A real-time intrusion detection technique was proposed in [3] using a supervised learning technique. A variety of methods were used to assess how well the IDS strategy performed and the decision tree fared better than other strategies, according to experimental findings. So, utilizing the decision tree technique, intrusion detection was developed to categorize the network data. In [4], a network generates and sends packets, which Wireshark then intercepts and examines for signs of intrusion. Using the Weka tool, the gathered data is preprocessed and then put up in a dataset with the chosen attributes. The data is then accurately classified using machine learning methods like naive Bayes, support vector machines, random forests, and K nearest neighbors with an accuracy of 83.63%, 98.23%, 99.81%, and 95.13%, respectively. The random forest has been found to be the best classifier with the highest level of accuracy for categorizing the packets.

The ensemble learning algorithm is a recent development in machine learning. It overcomes learning and increases the classification impact by merging the predictions from various base classifiers in accordance with a certain method or rule. The bagging algorithm is a classification technique used in [5] that makes use of integrated learning. T. V. Khoa [6] proposed a collaborative learning model for intrusion detection systems in the IoT. They created intelligent "filters" for this system, which are subsequently installed at IoT gateways to quickly identify and stop invaders. Each filter employs the network's accumulated data to develop a deep learning-based model for attack detection. The trained model will then be distributed to more IoT gateways to increase the system's overall accuracy in detecting intrusions. In the paper [7], the authors introduced the acceptable semantic technique for ensuring privacy in federated learning. Instead of using differential privacy-based systems, this strategy tries to increase data value and model efficiency while permitting a certain level of privacy that complies with privacy regulatory frameworks. In [8], the authors proposed a denial of service (DoS) detection attack at the victim side. In their system, the feed-forward neural network (FNN) is enhanced to detect DoS attacks reliably and efficiently. An ANN classifier is optimized to precisely detect DoS traffic and adopt unsupervised CFS technique for quickly and cheaply choosing important aspects of a DoS assault.

3 Methodology

In the methodology proposed by [1], they were able to reach the subset of 24 features. The accuracy, TPR, and FPR of the system are three key performance metrics that are taken into account while evaluating the usefulness of a feature. The critical set of features was found via sequential search. The classifier's accuracy was reduced by removing each characteristic one at a time till it reached a certain point. The experiment is performed 41 times to confirm that each feature in the NSL-KDD dataset is either significant, insignificant, or less significant. A 41-feature dataset is first submitted to a naïve Bayes classifier, and its output—which includes accuracy, RMSE, and the average TPR value—is then fed into the present method.

In [2], the ISCX Consortium generated the CICIDS-2017 dataset. The dataset is made up of eight traffic monitoring sessions, which are individually stored as CSV files (comma-separated values). The dataset contains two different categories of network: normal and anomalous (attacks). Some of the preprocessing processes were performed on the CICIDS-2017 raw dataset. This preprocessing produces a clean dataset. This study uses 30% of the data as the testing dataset and 70% of the data as the training dataset for the experiment.

In [3], using information gain as the feature selection criterion, 12 crucial characteristics of network data were discovered that are pertinent to identifying network attacks. The RT-IDS has a detection rate of more than 98% and can identify the two primary attacks within 2 s. To lower the FAR and boost the intrusion detection system's dependability and detection accuracy, a new post-processing method was devised. Each record after preprocessing is categorized as either ordinary data or attack data. A chosen machine learning method is trained as an IDS model in the first step. Each model is then tested using a fresh or new dataset where every row is recorded from the network. A majority vote mechanism is suggested. If the majority of records in a group are outlined to be from the same attack type, that group of records is regarded as the attack type. In any other case, the information is regarded as typical network activity. Through repeated iterations, this method ensures that no data point from the dataset is overlooked and is included in a cluster.

The methodology of [4] involves multiple steps. Firstly, a variety of packet kinds, including HTTP, HTTPS, and TLS, are generated. Using the dataset, the dataset's critical attributes that affect how packets are classified are chosen. Next, various machine learning techniques, including naive Bayes, random forest, support vector machines, and K nearest neighbors, are applied to the generated dataset to get the classification precisely. For training and testing purposes, the dataset is split in a ratio of 80% and 20%, respectively.

The methodology outlined in [5] makes use of an integrated learning classification system. Employing Bootstrap, the bagging method's fundamental premise is sampling from the training set in order to get sample sets at first. After that, from the basis classifiers' combination approach, a number of base classifiers trained on these sample sets are combined to produce an integrated classifier (see Fig. 1). The bagging approach provides the benefits of great generalization capacity, minimal model variation, and insensitivity to noise.

In [6], two machine learning-based approaches that can be applied in diverse IoT Industry 4.0 network scenarios were suggested. To specifically detect cyberattacks, they developed collaborative learning approaches based on classification and anomaly detection. For the attack detection system, the classification-based method can be utilized to forecast and determine the behavior of arriving packets.

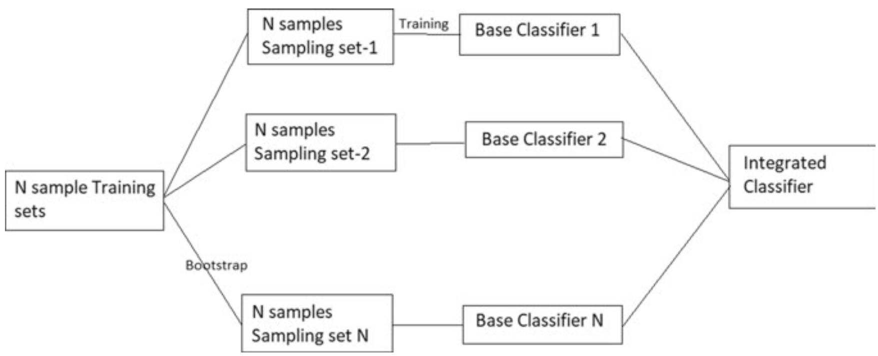


Fig. 1 Bagging algorithm flow

When the cyberattack detection system is only provided with an unlabeled dataset to build the neural network, anomaly detection-based collaborative learning [6] is beneficial in understanding anomalies.

Each site must choose the QID properties it should use to train its local model as the first step in [7]'s suggested methodology. In order to determine the most discriminative QID features and exclude all others from the local model's training, they precisely ranked the QID characteristics in accordance with the relevance of each feature. The next step is to choose a suitable syntactic strategy for each site's local data anonymization. Since unsolicited datasets were used to train each generative model, the global model's knowledge will be aggregated at a certain level.

Following FL convention, they used their syntactically anonymized data rather than their original data to train and share the global model among all sites. After the FL model is built, it can be used to forecast the data that is received at the server [7].

In [8], the suggested DoS detection method's basic framework is composed as:

Network traffic collector module: It is an application that the victim's edge network routers have installed. The victims' routers' inbound network packets are gathered by this module.

Data preprocessing module: It is responsible for standardizing feature values and choosing pertinent features for DoS detection. In the dataset, attribute values are typically not scattered evenly.

DoS detection module: It is in charge of the victim's routers' classification of incoming network data.

4 Result Analysis

The empirical results of [1] show that no available feature extraction method outperforms any other for intrusion detection. Even though the FVBRM produced reduced feature set which is the largest, it outperformed other approaches. In comparison with other methods, the FVBRM technique had the highest overall classifier accuracy (97.78%). Standard metrics including the confusion matrix, TPR, FPR, and accuracy were used to assess the results of the classifier.

According to experimental findings of [2], when checked with IDSs that use naive Bayesian (NB) and random forest (RF) classifiers, the suggested MAS-J48-IDS offers the best accuracy performance, with a 99.8% accuracy rate for all specified characteristics. Additionally, the utilization of MAS helps with load balancing within the network's nodes. The suggested MAS-J48-IDS performs better than random forest-based IDS but slightly worse than naive Bayesian-based IDS. The suggested IDS can regularly identify DoS/DDoS threats by utilizing a variety of attributes.

According to the experimental findings of [3], there were overall detection rates of 99.17%, 98.73%, and 99.43% for DoS assaults, probe attacks, and ordinary network activity. A sum of 0.07% of the DoS attack was incorrectly categorized as ordinary data, and 0.76% as a probe. The probe attack was mistaken for a DoS or normal data 1.27 percent of the time. The improper labelling of 0.08% of the regular network traffic as a DoS and 0.47% as a probe led to a false-alarm rate of 0.53%. A false alert occurs when normal data is misclassified as an assault. Each record took around 2 s to detect. Because network packets are aggregated every two seconds, this time is mostly used for data preparation, which takes about two seconds. Data classification takes only a few milliseconds. The post-processing approach further improves the normal detection rate to 99.979%.

Naive Bayes, random forest, K nearest neighbors, and support vector machine are the multiple classifiers that were used on the dataset of [4] for data prediction and packet classification into various categories. It is shown that the random forest algorithm outperforms the other machine learning algorithms used in this system for the prediction and classification of several types of data packets (Table 1).

The research proposed bagged tree-based intrusion detection model [5] has a good efficiency and a reduced rate of false alarms when compared to previous anomaly detection models that employ machine learning classifier approaches. The rate and proportion of missed alarms are also below average. In general, this model performs better in terms of detection. The approach described in this article [5] has an accuracy rate of more than 99% and provides an excellent detection impact for general attacks, DOS assaults, and other attacks.

The authors of [6] divided the dataset into T distinct subnetworks for the collaborative learning-based methods, including two subnetworks (Co-DL2) and three subnetworks (Co-DL3). While using the KDD dataset and Co-DL3, they were able to perform at their peak. The Co-DL2 has a similar pattern. Even though the Co-DL2's accuracy is 1.5% lower than the Co- DL3's, it can still surpass other traditional learning strategies.

To produce standard results supported by differential privacy, the authors looked at the trade-off between privacy and utility for a certain range of the privacy parameter in [7]. The level of privacy decreases as the privacy parameter ε rises, making the models more useful. All three classification methods agree on this. The level of performance decrease in FL, however, is substantially greater when using ε-differential privacy.

The UNSW-NB15 and NSL-KDD datasets are used in [8] to judge the effectiveness of the ADDM and u-MLP. The ADDM obtained maximum efficiency in the shortest amount of time: 99.2% on NSL-KDD and 97.1% on UNSW-NB15, respectively. In contrast, the u-MLP scored 79.2% on the UNSW-NB15 in 3.05 s and 83.5% on the NSL-KDD in 2.16 s. The remaining prediction accuracy rates are 98%, 81.2%, 92%, and 81.34% for DDMA, NSL-ANN, HSV-ANN, and ANN, respectively.

Table 1 Evaluation metrics of the algorithms used

SI. No.	Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
1	Naive Bayes	83.63	90.83	83.62	84.78
2	Random forest	99.81	99.81	99.81	99.76
3	Support vector machine	98.23	98.23	98.23	98.19
4	K nearest neighbors	95.13	94.54	95.13	94.46

5 Conclusion

From the above-mentioned information, it can be concluded that a lot of solutions came up for the most concerning problem of intrusion in a network. The development of ML-enabled IDS approaches is based on the analysis of internet traffic from devices to identify risk factors and assaults. Later, to prevent parties from sharing their data, FL, a collaborative learning approach, was developed in which an artificial intelligence algorithm is built locally across numerous decentralized endpoints, referred to as clients or parties. Through several training rounds, the data is continuously updated onto a global model. FL is a means of collaborative learning. In contrast to sharing their data, parties exchange their models with an aggregator, which evaluates a global model.

References

1. Mukherjee S, Sharma N et al, Intrusion detection using Naive Bayes classifier with feature reduction
2. Kerim B (2020) Securing IoT network against DDoS attacks using multi-agent IDS. In: 5th international conference on computing and applied informatics (ICCAI 2020)
3. Sangkatsanee P, Charnsripinyo C et al, Practical real-time intrusion detection using machine learning approaches
4. Thaseen S, Poorva B et al, Network intrusion detection using machine learning techniques
5. Chen P, Li F, Li J, et al, Research on intrusion detection model based on bagged tree
6. Khoa TV, Saputra YM, Hoang DT, Trung NL, Nguyen D, Ha NV, Dutkiewicz E (2020) Collaborative learning model for cyberattack detection systems in IoT industry 4.0. In: 2020 IEEE wireless communications and networking conference (WCNC). IEEE, pp 1–6
7. Choudhury O, Gkoulalas-Divanis A, Salonidis T, Sylla I, Park Y, Hsu G, Das A (2020) Anonymizing data for privacy-preserving federated learning. arXiv preprint arXiv:2002.09096
8. Idhammad M, Afdel K, Belouch M (2017) DoS detection method based on artificial neural networks. (IJACSA) Int J Adv Comput Sci Appl 8(4)

High-Resolution Image Inpainting Using Generative Adversarial Networks

G. Gopichand, A. Vijaya Krishna, I. Ravi Prakash Reddy, D. Vandana, K. Ramana, and P. Purshotham

Abstract In the twenty-first century, deep learning and artificial intelligence are being used to solve multiple problems in the real world. When coming to images, removing certain unwanted objects from images as per our interest has been one of the most famous intriguing problem statements in photo-editing. Image inpainting is a technique which is used to fill the missing regions present in images is a promising yet difficult activity. Previous approaches have accomplished critical advancement using GANs. However, these methodologies might experience the ill effects of reproducing distorted images and blurred textures in high-resolution pictures. The difficulties primarily come from content understanding from fine-grained and distant contexts image content understanding in a missing region in an image. Being an avid researcher and enthusiast in the field of artificial intelligence, working in such a problem statement would be a good way to hone the skills in the field. We propose an enhanced model using GANs for image inpainting of high-resolution images. The enhanced block helps in capturing both rich patterns as well as context from distant regions. The discriminator is also tweaked to make sure the robustness of the model.

Keywords Deep learning · Generative adversarial networks · Artificial intelligence · Context reasoning · Image inpainting

G. Gopichand
SCOPE, Vellore Institute of Technology, Vellore, Tamilnadu, India

A. Vijaya Krishna · P. Purshotham
Department of Computer Science and Technology, GNITS, Hyderabad, India

I. Ravi Prakash Reddy · D. Vandana (✉)
Department of Information Technology, GNITS, Hyderabad, India
e-mail: vandanadharmapuri1711@gnits.ac.in

K. Ramana
Department of Information Technology, CBIT, Hyderabad, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_73

1 Introduction

Image inpainting targets filling missing regions in images with reasonable contexts. A viable algorithm has a significant part in various applications, like photograph restoration, removal of objects and diminished reality. Regardless of the extraordinary advantages of this application, image inpainting faces huge hurdles in recovering sensible content and clear textures for large missing areas in high-resolution images. In [1], Liu mentioned that early works were able to help with the issue by either context based or patch-based algorithms. From [2], we could see that Deng stated that an algorithm based on diffusion spreads logical data from the boundaries into the empty region's direction. Patch-based algorithm performs the inpainting by copy pasting comparative patches from undisturbed contexts or external source. These methodologies can function effectively in completing thin holes in constant background. Generally, they regularly fail to generate appreciable content in semantic inpainting due to the shortage of efficient predicting of content that is missing and the textures present in complex scenes. Through a joint streamlining by adversarial losses and reconstruction losses, earlier models have given appreciable results in inpainting. It is said that straightforwardly using these models in images of higher resolutions have resulted in images with blurry textures and implausible structures. Such a problem makes these algorithms ineffective to use in practical applications, where users' photographs are regularly larger than a certain amount of resolution. Heusel, in [3], told that the difficulties fundamentally vary from trying to deducing sensible contents from image contexts and creating accurate image prediction for every pixel present in a wide region. The general 2 hurdles in this task are context reasoning and fine-grained texture synthesis. To beat the over two difficulties, we developed a novel model for image inpainting. Moreover, we tend to use a unique prediction task for the discriminator. In [4], Girshick told that maximum of the previous approaches followed PatchGAN for predicting the inpainted images as untrue, not paying attention to the fact that the missing regions outside actually come from real images, which make them struggle in generating realistic textures. To overcome this issue, we create a discriminator that could predict a down sampled mask. From Su's study from [5], we can understand this sorting of learning generally gives rise to a more efficient discriminator and may in turn help in improving the generator in synthesizing the realistic textures.

2 Related Study

In [6], Chen and group, the authors have designed dense up sampling convolution which is used generate pixel level predictions, which was able to capture more information than the more used bilinear sampling. They also proposed a hybrid convolution framework which could effectively enlarge the receptive fields which aggregates the global information and also reduced the so called "griddling issue" caused by the standard convolution operation and achieved 80.1% accuracy on the test set. Though this paper has stated about a model which has a high accuracy and an efficient way of inpainting the missing regions of an image, the work was only efficient towards images of lower resolution. When tested against a high-resolution image, the model gave a blurry output.

In [7], Lin's team have presented a generative inpainting system which functions on gated convolutions. This paper has been done on millions of images. This approach also solves the issue of vanilla convolution. They have presented a patch-based GAN loss called as SN-PatchGAN. Their work has produced higher quality and more flexible results than the previous researches at that time. The model that was developed in this research produced higher quality outputs, and the main drawback with this work was that certain sections of boundaries were not producing appropriate outputs. Also, their model was trained on millions of images, and it will be extremely difficult for other researchers to tweak and retrain their model and expect the same accuracy.

Jiang and team from [1] have proposed a deep learning-based approach with a new attention layer, which could not only preserve the contextual structure, but also make better predictions for the regions that are missing by learning the relevance between the patches in the missing regions. Unlike many other approaches, these authors have use U-Net architecture to approach the problem statement. They have used the most prominently used data sets as well. They have proposed a consistency loss to enforce the CSA layers in the decoder for them to be close to the VGG feature layer of the actual image simultaneously. Though this paper has stated about a model which has a better way of predicting the missing regions in an image, the work was only efficient towards images of lower resolution. When tested against a high-resolution image, the model gave a blurry output.

In [8], Yu has proposed an inpainting system that disentangles structure inference and content completion. They have also used a unique approach of predicting the foreground contour first, which helped them in making the process of inpainting more efficient. This method has made predicting of contours more effective, which substantially improved the performance of the inpainting model. Though this work had a unique approach to solve the problem, the problem arose when this model was used to remove natural contents such as spectacles and lamps. Moreover, this model was not effective in removing larger objects from the image.

Yeh's research team from [9] authors have proposed semantic image inpainting method by constraining on the data that is available for generating the missing content. Their model checks for the closest encoding of the distorted image using the context and prior losses developed by them. Unlike the state-of-the-art method present when the paper was published, their method performed the inpainting irrespective of the structure of the missing content. Though the model worked well irrespective of the structure of the missing content, the place where this model falls behind is when the size of the missing region comes into play. If the region that has to be edited is large and inappropriate, blurry and blocky images are obtained.

Yang's group from [10] have developed a model which could visually produce coherent completion for an image having incomplete regions. They have employed a new technique of dividing the problem into 2 steps and model them using a deep neural network. Unlike other works, they have used simpler methods to help the propagation of textures that are local which made the task of translation of image features and functions in a lesser space. This is the paper which was closest to what I want to achieve. But the area where this paper falls short is that even though it works on high resolution images, this model produces blurry outputs for certain colours and also takes a lot of time to do the processing of the image.

3 Proposed Approach

Generating clear textures and creating clear surfaces for an enormous freestyle incomplete region is an essential yet difficult to high-resolution picture inpainting. Here, we show a description of the developed network architecture which upgrades the texture synthesis and context reasoning for high-res picture inpainting. We propose a novel deep learning-based model for image inpainting of high-resolution images. The proposed network is able to gather the patterns of interest as well as the distant contexts. Moreover, we tend to use a unique prediction task for the discriminator. Maximum of the previous approaches followed PatchGAN for predicting the in painted photographs as untrue, not giving importance to the truth that the missing regions present actually come from real images, which make them struggle in generating plausible fine textures. To solve this issue, we created a discriminator which was able to predict a down sampled mask this sorting of learning generally gives rise to a more efficient discriminator which may further help in improving the generator in synthesizing plausible textures. Taking from what Yi mentioned in [11], we used a joint optimization of adversarial loss, style loss, perceptual loss and reconstruction loss. The generator has been modified with an encoder, a couple of ResNet blocks and a decoder. The input to the generator would be an image which will be masked and a mask depicting the masked regions and the output would be a completed image. Similar to general neural networks, the initial layers are stacked with convolution layers in the encoder to encode the input features. Secondly, these encoded features are passed onto the developed ResNet block, through which the generator captures informative contexts in images and patterns of interest for image inpainting.

Lastly, for decoding purposes, we used multiple layers of transposed convolutions. The developed block splits the kernel into different sub-kernels, each having lesser number of output channels. It also transforms the input features, by utilizing multiple dilation rates. This would spread out zeros between consecutive positions, which is very famous in semantic segmentation. In [12], Ren mentioned that a bigger rate makes the sub-kernel to view more area in an image, whereas the one with a lesser rate makes it view the local patterns in a field of smaller size. Transformations from various fields gets concatenated which is followed by a convolution layer for fusion of features. This permits the developed block to guess every (output) pixel through multiple views. From the prior steps, the developed block was able to perform various transformations which helps in increasing the context reasoning of the developed network. Moreover, it enriches the pathways of the combinations in the generator

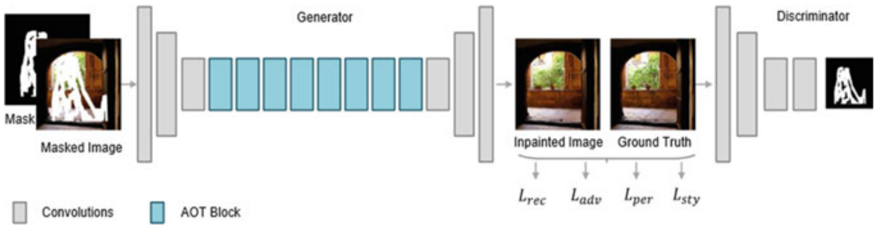


Fig. 1 High-level presentation of working of the model

network, which makes the generator to get as many patterns for context reasoning as possible.

We also developed a novel discriminator for the mask prediction task. As Bertalmio mentioned in [13], even synthesizing the textures for a large missing region is a difficult task as there will be multiple viable results for a missing region, wherein deep learning (inpainting) models are trained to fill the missing regions using reconstruction loss such as L1 loss. Gilbert in [14] told that generally, deep learning models produce a mean of the possible solutions, which would lead to blurry outputs. To overcome the difficulties stated by Aila from [15], we have used a joint optimization of adversarial loss and reconstructive losses, which helped us in generating clearer outputs. The main reason of choosing optimization for image inpainting is to make sure that we can achieve both visual fidelity as well as per-pixel reconstruction of inpainting images. Figure 1 shows a very high-level functioning of the paper.

3.1 Implementation Details

In building the generator, we stacked three convolution layers in the encoder layer, which is used to reduce the images' spatial space by 4 times. Then, three transposed convolution layers are used for upsampling in the decoder. Unlike the work did by Qin in [16], the developed ResNet block has been stacked in 8 layers as a main component in the generator. As seen in [17], Miyato suggested that when coming to the discriminator's arrangement, it is better to stack 4 convolution layers with stride value of 2 followed by PatchGAN. As Sun advised in [18], for the Gaussian filtering, the kernel size is set to 70 $\times$ 70 in order to stimulate the propagation of pixels in the convolution layer in the discriminator network. For the training of the model, 8 images were randomly sampled and masks were created. Referring to Zhu's work from [19], the learning rate we decided to use was 1e$-$4 for both discriminator and generator. As Murphy suggested in [20], the optimizer that was used was Adam with $\beta_2 = 0.9$ and $\beta_1 = 0$.

Fig. 2 Working of our model when compared to general residual block

3.2 Data Sets

As Guo suggested in [21], free form masks were used for both testing and training. All the images are resized to 512 × 512 following a standard setting. We have used 3 data sets for the training and testing purpose. The first one is the Places2 data set, which consists of 1,800,000 images from 365 different varieties of scenes. The second data set is the CELEBA-HQ, a high-quality data set consisting of human faces. This data set consists of 30,000 images, out of which 28,000 are used for training and 2000 are used for evaluation. The third data set is QMUL-OpenLogo which consists of logo classes. The reason of choosing these 3 data sets is that these 3 data sets are very different and very diverse. This would help the model to work accurately for almost any scenario (Fig. 2).

3.3 Real Applications

The proposed model can be used for multiple real-world applications such as removal of logos, editing of facial features and removal of objects from a photograph. We presented the results of logo removal in Fig. 4, where we could see that our model was very accurate with logo removal and was able to fill the images with plausible contexts. Coming to face editing, the results of our model have been presented in Fig. 3. The results show that our model could produce completely plausible structures of mouths and faces. Similarly, the proposed model also works quite effectively when it comes to object removal. The results depict that the developed model was effective in removing objects from various types of complex scenes as well as synthesize clear textures in an image. Figure 5 provides with examples of our model working on object removal.

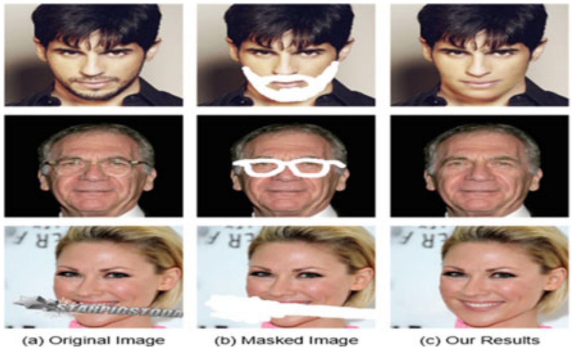


Fig. 3 Results of our model in facial editing

Fig. 4 Results of our model in the usage of logo removal from images

Fig. 5 Results of our model in the use case of object removal

4 Conclusion

In this paper, we proposed a novel model which could work effectively for image inpainting of high-resolution images. The model showed good results when compared to previous works. We also tweaked the discriminator using a novel approach, to make the generator synthesize much realistic textures. We have also provided figures stating the working of our model. As a future work, we have planned to extend our work to other problem statements, such as image denoising, image to image translation and single image super resolution.

References

1. Liu H, Jiang B, Xiao Y, Yang C (2019) Coherent semantic attention for image inpainting. In: ICCV, pp 4170–4179
2. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009) Imagenet: a large-scale hierarchical image database. In: CVPR, pp 248–255
3. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S (2017) Gans trained by a two time-scale update rule converge to a local nash equilibrium. In: NeurIPS, pp 6626–6637
4. Xie S, Girshick R, Dollar P, Tu Z, He K (2017) Aggregated residual transformations for deep neural networks. In: CVPR, pp 1492–1500
5. Su H, Zhu X, Gong S (2018) Open logo detection challenge. In: BMVC
6. Wang P, Chen P, Yuan Y, Liu D, Huang Z, Hou X, Cottrell G (2018) Understanding convolution for semantic segmentation. In: WACV, pp 1451–1460
7. Yu J, Lin Z, Yang J, Shen X, Lu X, Huang TS (2019) Free-form image inpainting with gated convolution. In: ICCV, pp 4471–4480
8. Xiong W, Yu J, Lin Z, Yang J, Lu X, Barnes C, Luo J (2019) Foreground-aware image inpainting. In: CVPR, pp 5840–5848
9. Yeh RA, Chen C, Yian Lim T, Schwing AG, Hasegawa-Johnson M, Do MN (2017) Semantic image inpainting with deep generative models. In: CVPR, pp 5485–5493
10. Song Y, Yang C, Lin Z, Liu X, Huang Q, Li H, Jay C (2018) Contextual-based image inpainting: infer, match, and translate. In: ECCV, pp 3–19
11. Yi Z, Tang Q, Azizi S, Jang D, Xu Z (2020) Contextual residual aggregation for ultra-high resolution image inpainting. In: CVPR, pp 7508–7517
12. Ren Y, Yu X, Zhang R, Li TH, Liu S, Li G (2019) Structure flow: image inpainting via structure-aware appearance flow. In: CVPR, pp 181–190
13. Bertalmio M, Vese L, Sapiro G, Osher S (2003) Simultaneous structure and texture image inpainting. TIP 12(8):882–889
14. Gilbert A, Trumble M, Hilton A, Collomosse J (2020) Inpainting of wide-baseline multiple viewpoint video. TVCG 26(7):2417–2428
15. Karras T, Aila T, Laine S, Lehtinen J (2018) Progressive growing of gans for improved quality, stability, and variation. In: ICLR
16. Qin C, He Z, Yao H, Cao F, Gao L (2018) Visible water mark removal scheme based on reversible data hiding and image inpainting. Signal Process Image Commun 60:160–172
17. Miyato T, Kataoka T, Koyama M, Yoshida Y (2018) Spectral normalization for generative adversarial networks. In: ICLR
18. Sun J, Yuan L, Jia J, Shum H-Y (2005) Image completion with structure propagation. TOG 24(3):861–868
19. Zhu X, Hu H, Lin S, Dai J (2019) Deformable convnets v2: more deformable, better results. In: CVPR, pp 9308–9316

20. Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2017) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS. TPAMI 40(4):834–848
21. Guo Q, Gao S, Zhang X, Yin Y, Zhang C (2018) Patch based image inpainting via two-stage low rank approximation. TVCG 24(6):2023–2036
22. Yu F, Koltun V (2016) Multi-scale context aggregation by dilated convolutions. In: ICLR

Intelligent Monitoring and Analyzer for Plant Systems

Umme Ayesha Samreen and B. V. Prasadbabu

Abstract In the growing busy life of people, taking care of plants has become a hectic task. In recent years, the need for automation is increasing day by day. Many automated require time-dependent motor activation control for watering the plants. In this project, a smart plant watering solution in the market is proposed. The purpose of this model is to predict which pesticides we need to add to the leaf of the plants. The camera helps in capturing the images of the diseased leaves. The ability to quickly identify diseases would be very beneficial because it will help in taking good care of plants at home.

Keywords Raspberry Pi3 B+ · DHT-22 · Soil moisture · Camera · CNN

1 Introduction

Nowadays, plants have a lot of benefits like they can enlighten the mood of depressed or sad people. The plants make the atmosphere happy, good, bright, and jolly in the home or farmhouse.

People in India are known popularly for growing plants in their homes or farmhouse. Therefore, yield and production can be boosted by combining different technologies present. This may raise production while simultaneously raising quality.

The most important component for diagnosing the illness of plants is the leaves. But if the disease is not detected and the plant comes in contact with the affected leaf, it affects the entire plant [1].

But as technology advances, people can use it and protect the plants. By doing this, we can protect the environment by feeding plants the right chemicals and preserving them as a result of lowering environmental pollution. Recognizing the right disease results in the proper use of chemicals, which lowers the cost and saves time and money [2–4]. In the past, we could do it manually by monitoring the leaf, but the

U. A. Samreen (✉) · B. V. Prasadbabu
G. Narayanamma Institute of Technology and Science, Hyderabad, India
e-mail: ayeshaumme.18@gmail.com

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_74

result was not always precise. We can now properly identify diseases in plants using cameras and deep learning algorithms, which is incredibly helpful. Technology is a good thing if we use it in a good way [5–7]. They are various new technologies that can help a lot in the maintenance of the plants like IoT and CNN. Internet of things and its various aspects like sensors (soil moisture, DHT-22, and camera) and microcontrollers (Raspberry Pi3 model B+) helps in detecting the atmosphere around the plants. CNN uses its various layers to analyze the images sent by the camera to the database and display the results to the user [8, 9].

The plant disease drastically reduces both the growth and yield of the plant. Therefore, it is necessary to find ways and means to combat plant diseases for maintaining the plants; if we design a good model, we can analyze and protect the plants before they affect the other plants in the surroundings [10, 11].

2 Literature Survey

The following data shows a review of various leaf disease detection and the techniques used to detect them.

1. The paper "Potato disease detection using machine learning" was written by Sadia Akter et al. and published in the year 2021; the techniques used are image processing and machine learning.
2. The paper "CNN-based disease detection approach on potato leaves" was written by Md. Asfaqur Rahman et al. and published in the year 2020; the techniques used are Open CV and CNN algorithm.
3. The paper "Maize leaf disease detection and classification using a machine learning algorithm" written by Himansu Das et al. was published in the year 2020, and the techniques used are machine learning and KNN.
4. The paper "Tomato plant leaf disease detection using CNN" was written by Sawan Rai and published in the year 2021, and the technique used is CNN.
5. The paper "Detection of Strawberry disease using CNN" was written by Hung-Yi Wu et al. and published in the year 2021, and the technique used is CNN.
6. The paper "Grape disease detection network based on multi-task learning and attention features" written by Somnath Dey et al. published in the year 2021, and the technique used is dual attention mechanism
7. The paper "Detection of pepper leaves diseases using CNN" and Machine Learning Algorithm written by Deepti Taterwal was published in the year 2021, and the techniques used are CNN and machine learning.
8. The paper "Detection and classification of Groundnut leaf diseases using KNN classifier" written by Vaishnnave M P et al. were published in the year 2019, and the technique used is a KNN classifier.

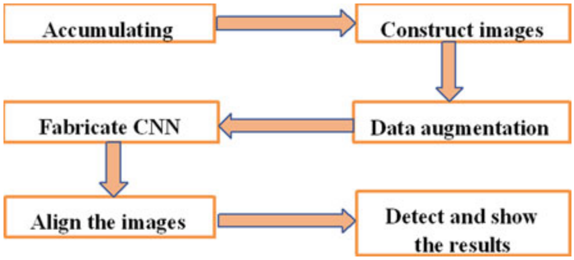


Fig. 1 Existing method block diagram

3 Existing Method

The identification of disease in potato leaf requires the steps shown in the block diagram (Fig. 1).

The images are captured and then they are created as a database/dataset which is trained and tested by CNN, and a model is created. This model identifies the result and displays it [3, 12].

4 Proposed Method

In this proposed system, we use a convolution neural network that helps to create a model which helps in training and testing to detect the disease of the leaf and predict the cure of the affected leaf of the plant. Here, we capture the images using a camera that can be placed near the plant. After the completion of the model, we can combine CNN with IoT. Here, the use of IoT is done, so that we can keep a complete eye on the plants and we use various sensors (DHT-22, soil moisture sensor), a camera, and various aspects of IoT to preserve the plants from getting damaged. The CNN and IoT together help in curing the disease detected by the above model before the whole plant is affected or before it spreads the disease to the other plants placed along with the affected plant. CNN uses its various layers to train and test the model which detects the disease and the cure (Fig. 2).

This system uses different sensors like soil moisture, DHT-22, and a camera. First comes the soil moisture sensor placed in the soil, so that it can detect the moisture in the soil whenever the moisture goes below the expected level, and then the plant can be watered according to the requirement of the plant. Next comes the DHT-22 sensor which is placed near the plant where it detects the humidity and temperature of the surrounding area, if any one of the temperatures or humidity is increased or both of them are increased then the user can water the plant placed in the home or farmhouse. Last, we have the camera which is placed near or above the plant from where it can capture the image of the spoilt or damaged, or affected leaf, and it sends the image to the database where the CNN algorithm uses its various features to detect the disease and predict the pesticides required to cure the disease.

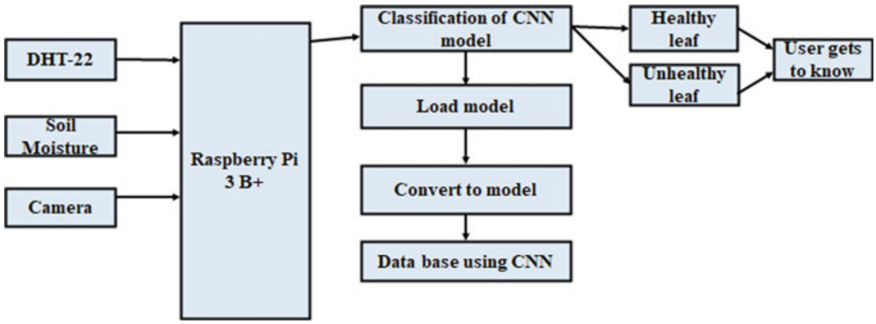


Fig. 2 Proposed system

4.1 Raspberry Pi3 B+

The image (Fig. 3) shows that it is a version 3 B+ model in the Raspberry Pi family. It has forty GPIO pins, four USBs, one audio, and one video port. The camera is connected to the microcontroller through the USB port.

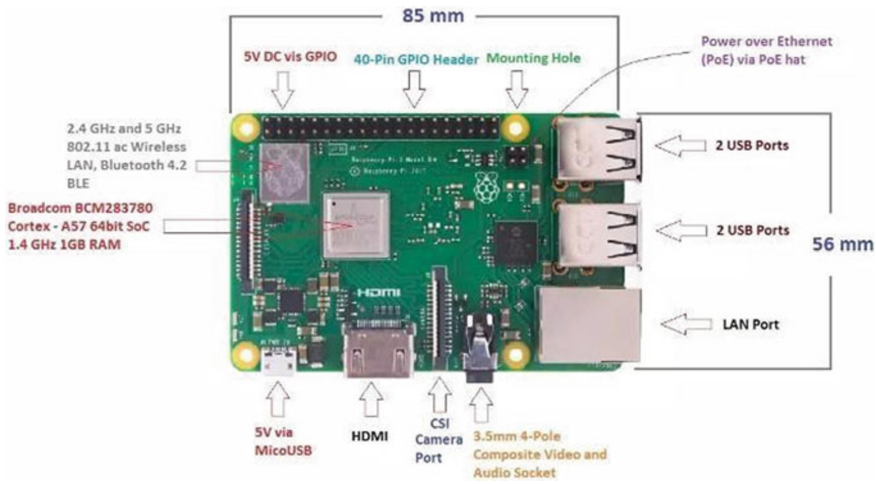


Fig. 3 Raspberry Pi3 B+

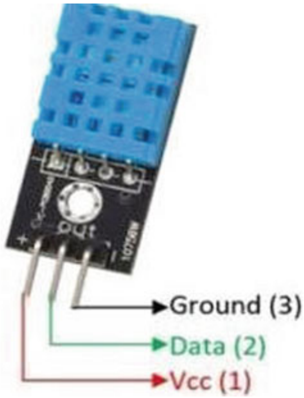


Fig. 4 DHT-22

4.2 DHT-22

In this model, the sensor is connected with the help of jumper wires. The VCC, GND, and the DOUT of this sensor are connected to VCC, GND, and GPIO pin, respectively, with the microcontroller. This sensor detects the temperature and humidity of the plant's surroundings (Fig. 4).

4.3 Soil Moisture

In this model, the sensor is connected with the help of jumper wires. The VCC, GND, and the DOUT of this sensor are connected to VCC, GND, and GPIO pin, respectively, with the microcontroller. This sensor detects the moisture in the soil near the plant (Fig. 5).

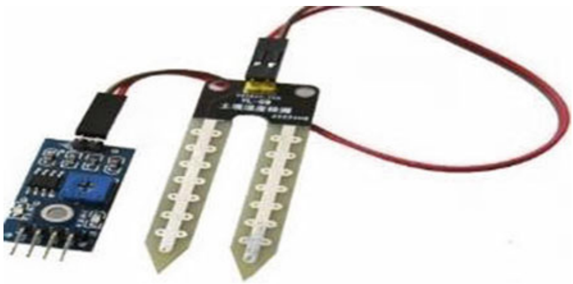


Fig. 5 Soil moisture

Fig. 6 Camera

4.4 Camera

The camera is connected to the microcontroller through the USB port in this model. The camera captures images of the diseased leaf (Fig. 6).

4.5 Jumper Wires

The wires which are used to connect the sensors and the Raspberry Pi are called jumper wires which are of three types which are male to male, female to female, and male to female.

4.6 Convolution Neural Network

The convolutional neural network is used to locate and classify objects in an image (Fig. 7).

CNN has several hidden layers which help extract data from images. The main tiers of CNN are as follows.

4.6.1 Convolution Layer

The process of removing useful components from an image begins with this.

4.6.2 ReLU

After being retrieved, the feature maps must next be transferred to a ReLU layer.

4.6.3 Pooling Layer

Now, a corrected feature map is run through a feature map with a pooled layer created (Figs. 8 and 9).

The convolutional layer receives the image's pixels and conducts the convolution, resulting in a convolved map.

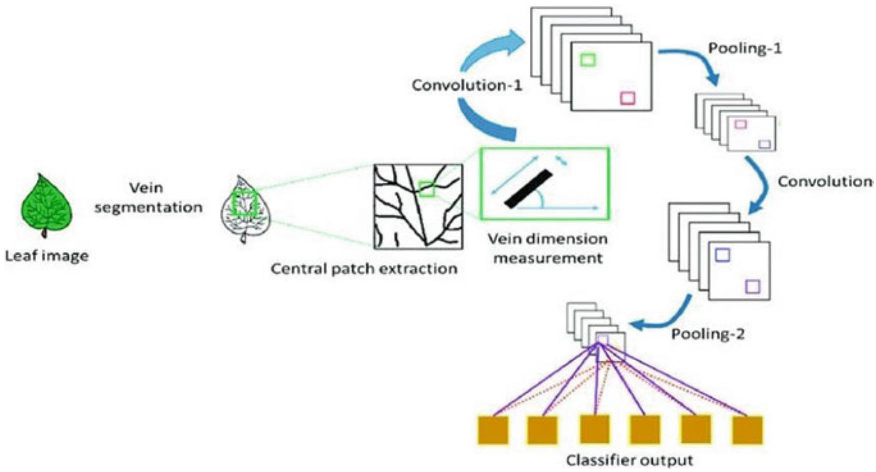


Fig. 7 CNN algorithm

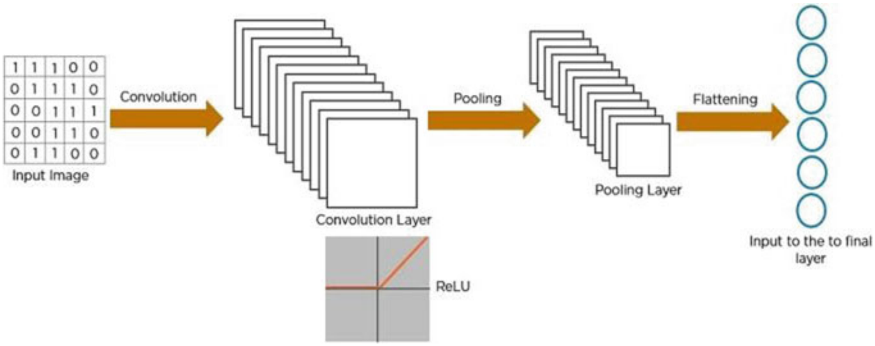


Fig. 8 Pooling layer

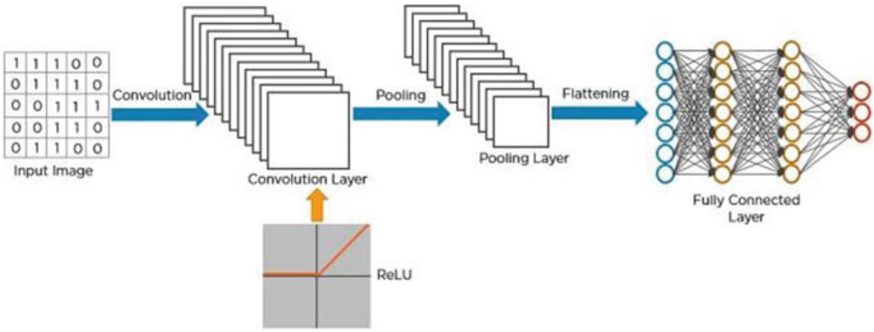


Fig. 9 Detailed image of pooling layer

4.6.4 Fully Connected Layer

Neurons in this layer have full connectivity with all neurons in the preceding and succeeding layers.

4.7 Dataset

The datasets/database used in this model is collected from sites like Kaggle, Google, and real life.

5 Result and Discussions

The designed CNN model is trained and tested with approximately 1000 images captured in real life and the images obtained from different sources like Kaggle of healthy and unhealthy leaves. To run the program, we open the Python idle and write the code and run the program, and a frame opens which captures the image of the leaf and it sends the image to the dataset/database where the image is compared with the images in the dataset/database, and the result is displayed if the leaf is affected with which disease and the pesticide to cure the disease. Along with the disease and the cure, the moisture of the soil and the humidity and temperature around the plant is also displayed to the user.

Below are a few examples of the expected result of this proposed system (Figs. 10, 11, 12 and 13).

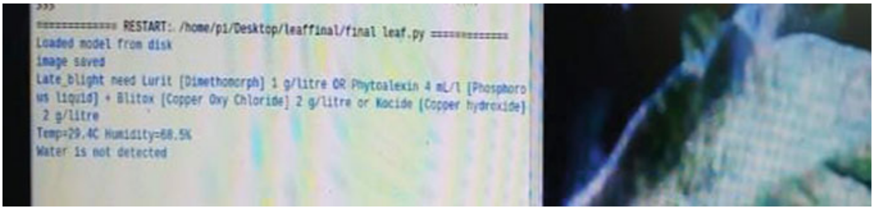


Fig. 10 Disease on the leaf is late blight and the cure for the disease

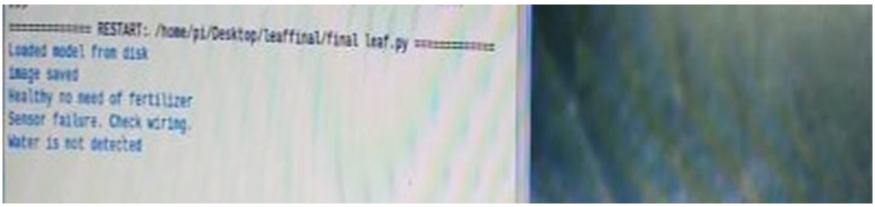


Fig. 11 Healthy and requires no cure

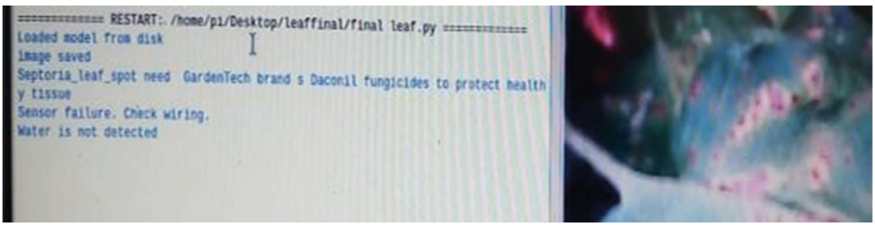


Fig. 12 Disease on the leaf is Septoria spot and the cure for the disease

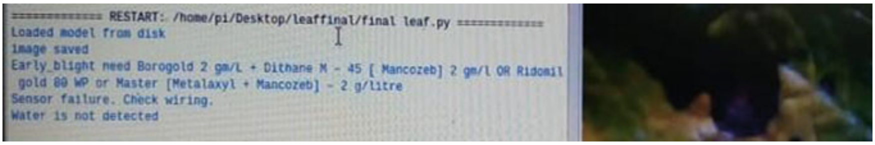


Fig. 13 Disease on the leaf is early blight and the cure for the disease

6 Conclusion and Future Scope

The advancement in the technologies like IoT, CNN, etc., has helped a lot in various ways in the detection of diseases on the leaf or the plant.

This paper concludes that the plants which are affected by the disease will be detected on time or in other words before it is spread to other parts or other leaves of the plants.

In the future, we can also predict if they are any insects present and which insecticides to be used and also add a user interface that will help the user in accessing the data easily from anywhere and anytime.

This technology can also be used in places where the plants can be grown in hydroponics conditions.

References

1. Anand R, Veni S, Aravinth J (2016) An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method. In: 2016 International conference on recent trends in information technology. IEEE
2. Sardogan M, Tuncer A, Ozen Y (2018) Plant leaf disease detection and classification based on CNN with LVQ algorithm. In: 2018 3rd International conference on computer science and engineering. IEEE
3. Asif MKR, Rahman MA et al (2020) CNN-based disease detection approach on potato leaves. In: 2020 3rd International conference on intelligent sustainable systems. IEEE
4. Francis J, Anoop BK (2016) Identification of leaf diseases in pepper plants using soft computing techniques. In: 2016 Conference on emerging devices and smart systems. IEEE.
5. Sharath DM et al (2019) Image-based plant disease detection in pomegranate plant for bacterial blight. In: 2019 International conference on communication and signal processing. IEEE
6. Vaishnnave MP et al (2019) Detection and classification of groundnut leaf diseases using KNN classifier. In: 2019 IEEE international conference on system, computation, automation, and networking. IEEE
7. Das D et al (2020) Leaf disease detection using support vector machine. In: 2020 International conference on communication and signal processing. IEEE
8. Padol PB, Yadav AA (2016) SVM classifier based grape leaf disease detection. In: 2016 Conference on advances in signal processing. IEEE
9. Sankar M, Mudgal DN, Jalinder MM (2019) Green leaf disease detection using raspberry pi. In: 2019 1st International conference on innovations in information and communication technology. IEEE
10. Dhingra G, Kumar V, Joshi HD (2018) Study of digital image processing techniques for leaf disease detection and classification. Multimed Tools Appl

11. Overbeek MV, Kaesmetan YR, Tobing FAT (2019) Identification of maize leaf diseases cause by fungus with digital image processing. In: 2019 5th International conference on new media studies. IEEE
12. Azlahhnique MAF et al (2019) Review on techniques for plant leaf classification and recognition

Intelligent Alert System for Women Security

G. Roja, B. Spandana, N. Divya, and Y. Chittibabu

Abstract Today in India, more than ever there is a great need for securing the freedom and protect the women to exercise their fundamental right of liberty and freedom. In spite of several acts promulgated against the atrocities taking place in the society. In order to safeguard and protect women from diabolic activities and to minimize the crime, we came up with this project INTELLIGENT ALERT SYSTEM FOR WOMEN SECURITY. These systems based on machine learning can be used for understanding the dangerous situations better than any other currently existing technology. With the implementation and use of this technology, it is possible to provide the level of freedom, care and help towards the women and feel comfortable even at nights, letting them work as equals to men and enjoy the freedom that they're supposed to have since the beginning. This system will provide a chance to live a stress free life by assuring them their security.

Keywords Women security · CNN · Convolution · Confusion matrix · Convoluted image matrix

G. Roja (✉) · B. Spandana · N. Divya
Department of Computer Science and Engineering, G. Narayanamma Institute of Technology and Science, Hyderabad, Telangana, India
e-mail: rojagurrapu@gnits.ac.in

B. Spandana
e-mail: spandana.d@gnits.ac.in

N. Divya
e-mail: n.divya38@gnits.ac.in

Y. Chittibabu
Department of Electronics and Communication Engineering, CMR College of Engineering and Technology, Hyderabad, Telangana, India

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_75

1 Introduction

A lady is the sign for kindness, freshness, understanding and abandonment. Harmony and success lie with inside the civilization in which a girl is happy and admired. Think of the female deity you bow right all the way down to with inside the equal residence wherein you improve your hand on the real goddess of your home. Nowadays, ladies are preserving tempo with guys in life, regrettably at fee of being subjected to abuse, harassment and violence in public or even at their personal houses. They can't step out in their homes at any time, they can't put on garments as in line with their wish, nor can they even move for paintings in peace. This all takes away their freedom however additionally unfastened their self-belief and dreams. According to National Crime Records Bureau, crime towards ladies has notably accelerated in current years. Due to the above reasons, it is miles pretty obvious that there is a striving want for ladies safety [1] within side the country. In past decades, women won't step out from their houses to work, so there was more safety. But in present scenario, women want to be employed, and they want to work outside. But there is lack of safety. One of the 1/3 of the girls might also additionally be afflicted by violence in her lifetime. Such incidents are extra not unusual place nowadays. There are many systems that have been built to provide safety foe women. There cannot be cop who always guarding the women, but there can be safety measures which women can used by them.

The proposed system is based on neural networks (CNNs) incorporated with an alerting system. This system will monitor the chosen place at regular intervals through a camera. Once any suspicious activity is detected, then it is going to ship statistics to the close by police station and to the encompassing stores and complexes through an e-mail alerting them alongside the vicinity and image captured.

2 Existing Systems

1. **"Self-protection machine for ladies with place monitoring and SMS alerting thru GSM network"**: This machine incorporates a surprise mechanism to supply non-deadly electric powered surprise in emergency conditions to discourage the attacker as quickly because the cause key at the band is pressed.
2. **"Safe: a girls protection machine"**: This machine includes 3 major additives particularly an android utility, major tool and transportable digital digicam [2]. Collectively, this gadgets will paint as an powerful protection machine major tool which include raspberry pi incorporated GPS guard alongside manually perform pepper spray. Android utility may be utilized in feasible methods, wherein it'll both use telecell smart phone GPS machine or it'll use GPS machine of major tool to seize location.
3. **"Depiction and improvement of an IoT primarily constructed totally attire tool for the protection and protection of girls and woman child"**: The intention of this paintings is to broaden a wearable tool for the protection and safety of

girls and girls. This goal is performed via way of means of the evaluation of physiological alerts alongside frame position [3]. The physiological alerts which might be analysed are galvanic pores and skin resistance and frame temperature. Real-time tracking of statistics is performed via way of means of wirelessly sending sensor statistics to an open supply Cloud Platform.

4. **"Intelligent protection gadget for girls security"**: They describe each time we experience dangerous she can be able to press the button of device, that occasion could be identified byLPC2148 controller [4]. Then controller will generate manage sign for GPS gadget, it is going to ship manage sign via MAX 232 to GPS module.GPS gets activated, so it is going to tune the precise vicinity of the sufferer and ship this records again ARM controller via MAX 232 interface.
5. **"Self-Salvation—the Women's Security Module"**: If any individual desires to make an strive on girls, then with the aid of using urgent a button, it'll ship alert message to the mother and father or pals telecell smart phone numbers and additionally to the closest police station. The technique of protection is thru GPRS we will song the girls/automobile position. With the aid of using sending SMS [5, 6] "TRACK" to the prevailing telecell smart phone wide variety, we will get the picture.

3 Proposed System

The proposed intelligent alert system for women security model is based on neural networks (NNs) incorporated with an alerting system. These systems based on machine learning can be used for understanding the dangerous situations better than any other currently existing technology. With the implementation and use of this technology, it is possible to provide the level of freedom, care and help towards the women and feel comfortable even at nights, letting them work as equals to men and enjoy the freedom that they're supposed to have since the beginning. This system will provide a chance to live a stress free life by assuring them their security. This system will monitor the chosen place at regular intervals through a camera. Once any suspicious activity is detected, then it is going to ship statistics to the close by police station [7] and to the encompassing stores and complexes with the aid of using an electronic mail alerting them in conjunction with the place and photograph captured. We train the machine using CNN model and use OpenCV [8] for process the video to detect any suspicious activity.

3.1 Methodology

A convolutional neural community (CNN or ConvNet) is a community structure for deep gaining knowledge of which learns immediately from information, getting rid of the desire for escort quality removal. CNN [9, 10] is mainly advantageous for identifying styles in pix to apprehend objects, features and sequences. They also can be pretty powerful for classifying non-picture information which include audio, time series and sign information.

Why we have chosen CNN:

1. CNN removes the desire for escort quality removal—the abilities are located at once via way of means of the CNN.
2. CNN fabricates notably veracious reputation results.
3. CNN perhaps re-educated for ultra-modern popularity tasks, allowing you to construct on pre-present networks. Four foremost operations with inside the CNN are proven as follows:

(i) **Convolution**: The foremost use of this operation with inside the situation of a CNN is to perceive suitable functions through the picture can be used as an entry point to the primary layer. This keeps the geometric affiliation of the pixel dots. Every picture be regarded as a mathematical N × N matrix of pixel dots every dot represents its very own value. Pixel dot be the smallest identifier on this picture mathematical N × N matrix.

Assume a five via way of means of five (five * five) mathematical 5 × 5 matrix having values are showed in duplex (i.e. zero or one), for high-level comprehension. The matrix dots are commonly having RGB values within the range of zero and 255 in total 256 pixel dots. The given diagram represents this will be elaborated now to allow us to taking a case of any other three via way of means of three matrix as represented as: So the convoluted picture matrix [11] could be as visible in Fig. 1.

Now allow us to take a case of every other three with the aid of using three matrices as represented as:

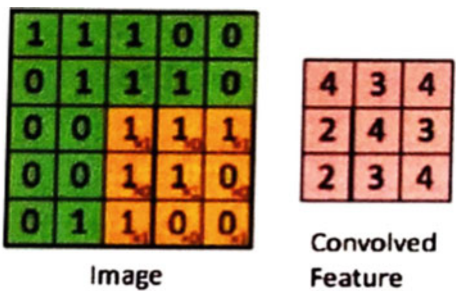


Fig. 1 Convoluted image matrix

1	0	1
0	1	0
1	0	1

So the convoluted photograph matrix could be.

(ii) **Inconsistent**: Corrected straight line unit is inconsistent operation. It applies as on a simple level. In different approach, it's a procedure to this carried out consistent with dots and supplants every of its non-advantageous values of every dot within the characteristic matrix through zero. This is essentially an easy conjecture Equation: $(1 + \text{ex}\,(x) = \text{Inf})$.

(iii) **Joining or multi-stage illustration**: Geometric joining will be likewise referred to as sub-sampling or down sampling which facilitates in decreasing the size of every function map however even at the same time as doing so, keeps the maximum consequential records of the map as visible in Fig. 2. After joining be done, subsequently their 3D function matrix is transformed into the only proportional function direction.

(iv) **Stratification Layer**: The result from the entanglement and joining operations affords outstanding functions can be brought out from the photograph as visible in Fig. 3. These functions afterwards used by the totally linked layer for lower the enter photograph converts extraordinary lessons premise at the education data set.

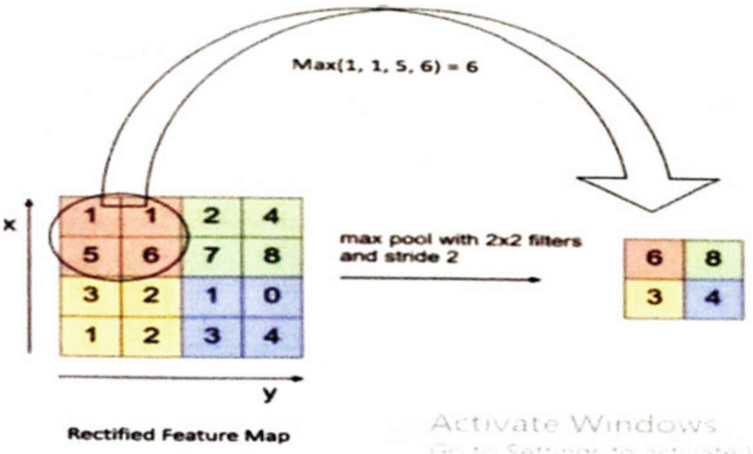


Fig. 2 Pooling and sub-pooling

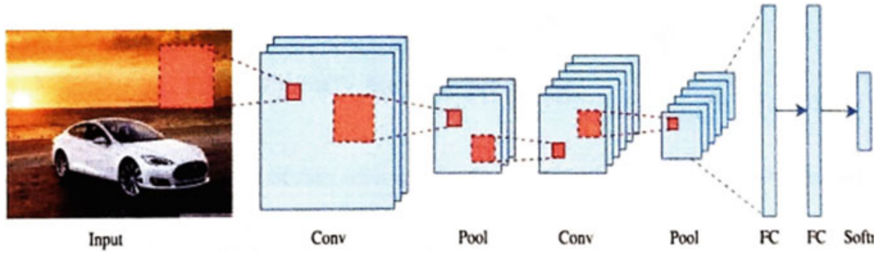


Fig. 3 Stratification (fully connected layer)

Fig. 4 OpenCV for video processing

Why we've got selected OpenCV

OpenCV Library is a global supply pc insightful and predictive and device getting to know s/w coding library.

1. OpenCV became constructed for offer a non-unusual place framework for pc insightful and predictive packages and to gear up the usage of device notion in business items.
2. OpenCV is an open supply library which gives us with the equipment to carry out nearly any type of photograph and video processing as proven in Fig. 4.
3. Using OpenCV (Python) would be very easy [12].

3.2 Objectives of the Project

Identification of criminal activities against women.

Immediate response, i.e. Find the fast help.

Protection of a person.

3.3 Architecture

See Fig. 5.

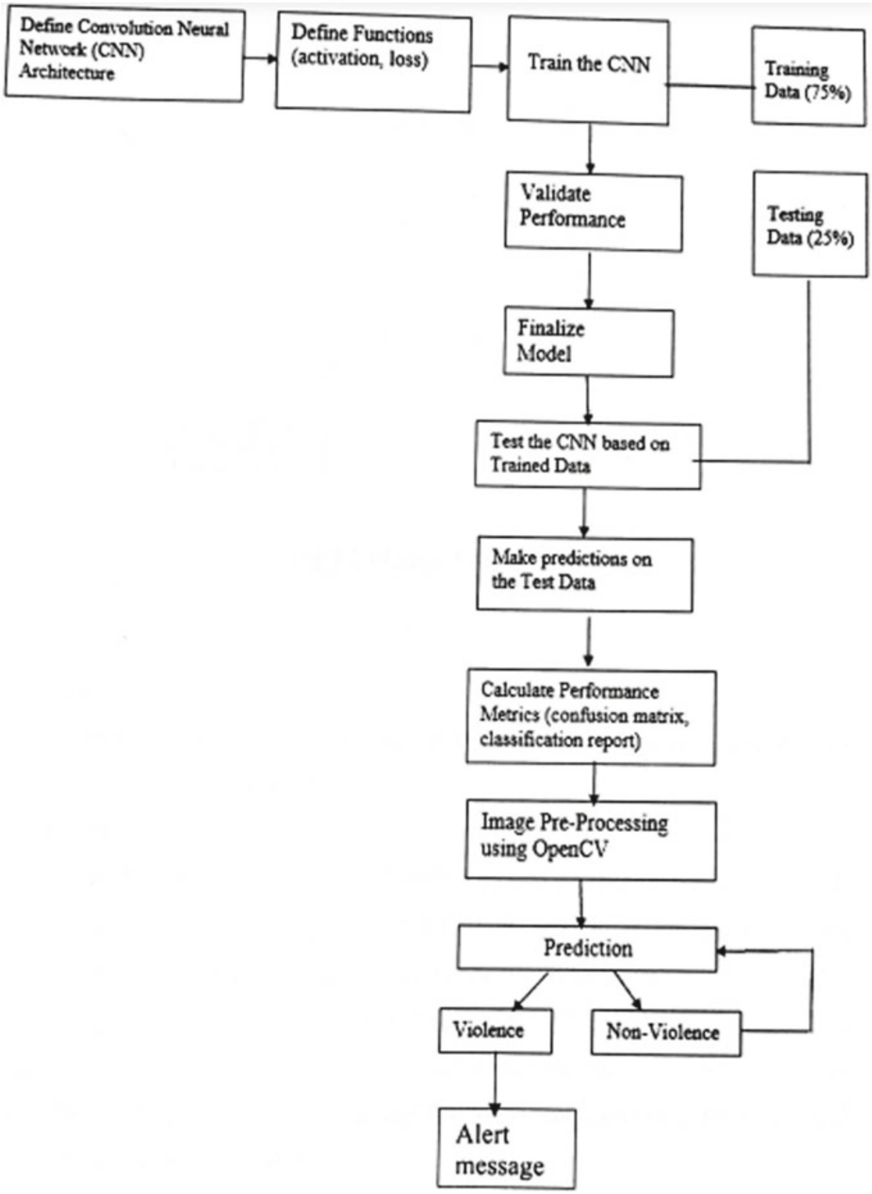


Fig. 5 Architecture of intelligent alert system for women system

3.4 Module Design

See Fig. 6.

Input Image

The enter photograph can be captured via way of means of the surveillance camera. This photograph can be processed via way of means of the OpenCV algorithm [13].

Pre-processing

Achieving excessive overall performance of violence detection structures calls for overcoming a few fundamental difficulties. Such as developing a database and unifying picture dimensions. In the subsequent section, the approach utilized in picture resizing is explained. Image Resizing: To clear up the trouble of various picture sizes with inside the database [14], an enter picture is both multiplied or reduced in length. Unifying the picture length gets the identical quantity of functions from all images. Moreover, resizing the picture reduces processing time and as a result will increase gadget overall performance.

Aspect Uprooting

Aspect uprooting is a manner of capacity reduction through which a preceding group of raw data is lowered to additional conceivable businesses for filtering. A function of these big information units will be a big quantity of dynamic values those need a different type of calculating benefits to manner. Aspect uprooting is the call for techniques it chooses and/or integrates dynamic values into aspects, correctly lowering the

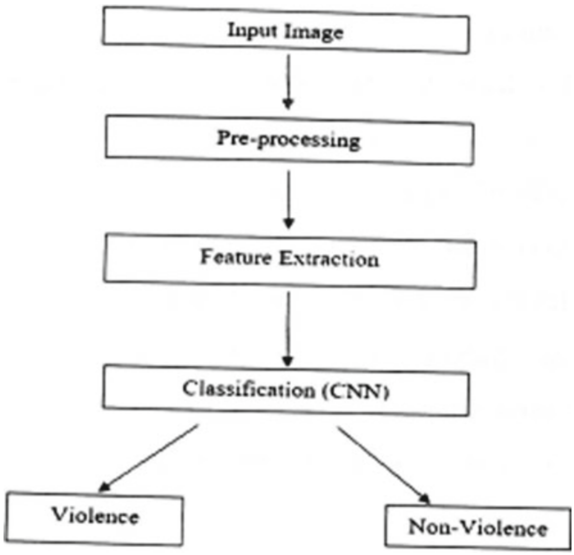


Fig. 6 Module design

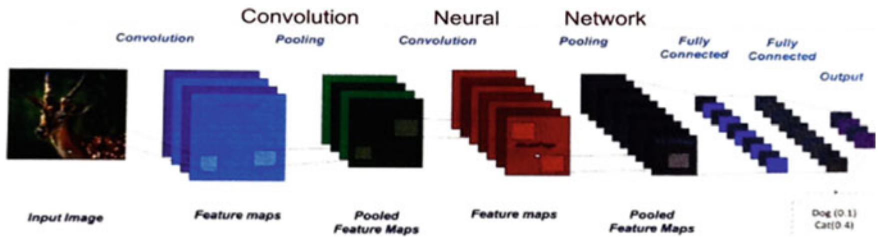


Fig. 7 Feature extraction

quantity of information that have to be cooked, whilst nevertheless correctly and absolutely explaining the authentic information group. The manner of aspect uprooting has advantage of whilst which we need to lower the quantity of assets wished for filtering without lowering crucial or important information. Aspect uprooting also lowers the size of repeated information for a particular analysis. Also, the slashed of the information and the machines endeavours in building dynamic values combinations (features) eases the rate of procured apprehension of and conception steps within side the device procured apprehension of manner.

Classification

Classification in device gaining knowledge of is whilst a device or pc makes use of an set of rules to attract conclusions from records that it already has, after which makes use of those conclusions to categorize new records it receives. Classification is a pc imaginative and prescient method. After extracting features, the position of category is to categorize the photograph through convolutional neural network (CNN) (Fig. 7).

4 Implementation of the Modules

4.1 Data Sets Use

4.1.1 Data Preparation

Context

The record includes snap shots of violence and non-violence taken from https://www.kaggle.com. The general wide variety of snap shots is round 14,442, out of which about 11,063 were break up with inside the schooling set and 3379 with inside the check set.

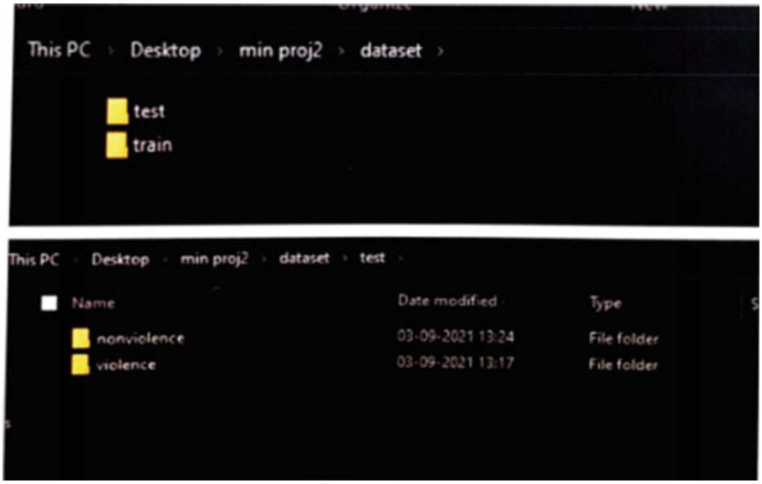


Fig. 8 Data sets

Content

The snap shots are in JPEG format, such as three channels, i.e. RGB. The resolutions range from photograph to photograph, and from class to class, however typical those aren;t extraordinarily high-decision imagery. The classes encompass violent and non-violent.

Acknowledgements

The pictures are taken from the general public portal Kaggle (https://www.kaggle.com.) That is the biggest data set supply online constructed for the motive of supplying distinct sorts of data sets (Fig. 8).

4.1.2 Training

The device may be widely classified into the subsequent fundamental phases:

1. **Image Acquisition**: The pictures are received both via a digital digicam and via a regionally saved device. Whatever is probably the resource, and it's far huge vital the enter photograph is obvious and accurate. To do this, a high-level pleasant photograph is needed.
2. **Photograph Pre-processing**: The picture is normalized on this segment through disposing of unwanted extra noise for example hair and pores and skin stains, because it may be able to distract the analysis. However, the picture that is

provided for the process will not be of trendy size as needed for the algorithm, because of this the vital specified picture size is obtained.

3. **Information base problem to preserve checking out and education records photographs**: In the sample of administered instruction, because the sample here, an education data group is essential. The checking out data group is the photographs obtained in the course of picture acquisition.

Activation Function Types

Twisted Function: Twisted function is a feature this is arranged as an "S" fashioned chart. Equation: $A = 1/(1 + e - x)$.
Type: Irregular, observe that X values ranges between $-$ 2 and 2 and Y values are very abrupt. It approaches very less adjustments to X which could additionally result in massive adjustments with inside the price for Y value.
***Y* Range**: zero to one.
Advantages: Generally used with inside the o/p layer of duplex categorization, in which the end result is both zero or 1, because the price for the sigmoid feature lies among zero and 1 most effective so, the end result may be expected effortlessly to be 1 if the price is extra than 0.5 and zero in any other case.

Optimization Techniques

Stochastic gradient descent (SGD) represents a version of rise. Rather of acting calculations at the entire data group—that is unnecessary and amateurish. SGD handiest calculates a less batch of a dynamic choice of information choices. SGD generates the equal overall presentation as ordinary rise whilst the mastering fee is less. Nowadays, a range of latest enhancers were suggested to address complicated schooling situations, wherein rise techniques work poorly. For example, maximum broadly included and sensible enhancers for schooling deep mastering fashions are Adam.

Adam is a set of rules for rise-primarily hinged totally minimization of problematic goal functions. It joins the benefits of SGD annexes. Adaptive Gradient (AdaGrad) algorithm and Root Mean Square Propagation (RMSProp) calculate person adjustable mastering quotes for distinctive values.

RMSprop stands for Root Mean Square Propagation. Widely acknowledged gradient descent optimization set of rules for mini-batch mastering of neural networks. It makes use of a shifting common of squared rises to synthesize the rise itself that has the impact of equalizing the step length—lowers the step for the big rise to keep away from blasting, and growth the step to the minor rise to keep away from disappearing.

Mixed destruction: Mixed destruction loss, or log loss, measures the overall productivity of a class version which result will be possibility fee among zero and one. Mixed destruction loss will increase because the expected possibility separate from the real label. The data set includes 3 categories (bacterial pneumonia, viral pneumonia, normal), the version is constructed the usage of express mixed destruction. Categorical mixed destruction is a loss feature, and this is utilized in multi-elegance

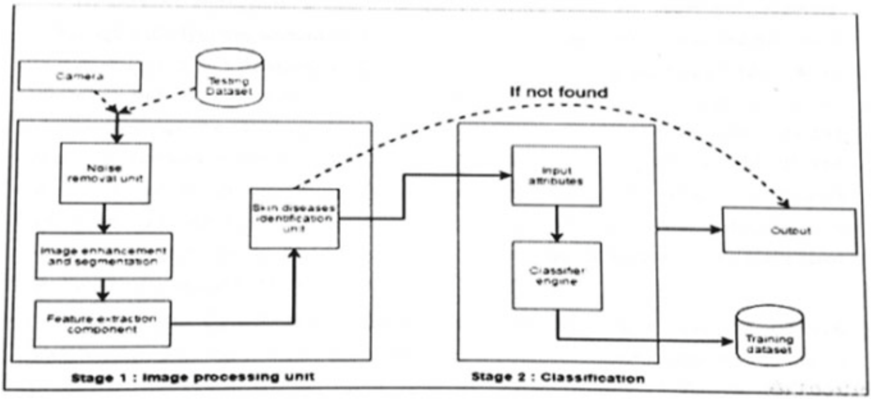


Fig. 9 Classifier to identify the violence

class duties. They are duties wherein an instance may simplest be owned by at least only one viable categories, and the version should determine to one.

Classifier to pick out the violence: Sigmoid organizer used right here will be the remaining surface of the community that submit a real opportunity of every tag (Fig. 9).

The structure includes essential parts: The photograph processing unit and the category unit. The photograph processing unit will beautify the photograph through eliminating nostril after which the photograph may be segmented into extraordinary segments to distinguish the functions of the photograph in an effort to be extracted to discover whether or not the image is violent or non-violent.

- Noise elimination method: Reduces the undesirable pixels from the image.
- Aspect Uprooting Component: Aspect uprooting is the foremost steps in any categorization-orientated problem. Aspects are the important thing for each education and checking out reasons. This function consists of critical statistics approximately the picture as a way to be used to become aware of the disease.
- Violence Recognition method: Identifies out either the picture is violent or not.
- Entered Parameters: All the critical parameters along with asymmetry, corner, colour, length, progress and so on where have been uprooted from the picture at the moment are given as enter to Part IT this is the organization part.
- Organizer Machine: Organizes the photos to the one of the prior-described lessons thru category algorithm (right here sigmoid classifier is used)

4.2 Technologies Used

Python: Python [12] is an interpreter, object-oriented, high-level programming code with different semantics. Its high-level integrated records arrangements, assorted with high-powered typing and high-powered binding, make it very engaging for robust application development, is also to be used as a scripting or adhesive language to attach present additives together. It is simple, easily verified syntax highlights intelligibility and therefore lowers the value of application continuity. It helps programmes and collections, which influence application extensibility and code reuse. This interpreter and the giant preferred codebase to be had in contributes or duplex shape without fee for all essential structures and may be freely distributed. We have used Python for reporting prediction of violence.

Machine Learning: Machine studying called to be a sort of synthetic intelligence (Al) that used in software programmes to emerge as greater true at analysing effects without having extraordinary programmed values to do so. Machine studying logics use historic facts as insert to anticipating new resulting values.

4.3 Algorithm

See Fig. 10.

5 Results and Discussions

5.1 Performance Measures

Confusion matrix: A confusion matrix is a way for concluding the overall production of class logic. Classification precision on own may be deceptive when we have a wrong variety of findings in every single magnificence or when we have larger than instructions for your data groups (Fig. 11).

True Positives (TPs): TPs are the instances whilst the true magnificence of the records factor turned into digit one (True) and the anticipated is likewise digit one (True).

True Negatives (TNs): TNs are the instances whilst the true magnificence of the records factor turned into digit zero (False) and the anticipated is likewise digit zero (False).

False Positives (FPs): FPs are the instances whilst the true magnificence of the records factor turned into digit zero (False) and the anticipated is likewise digit one

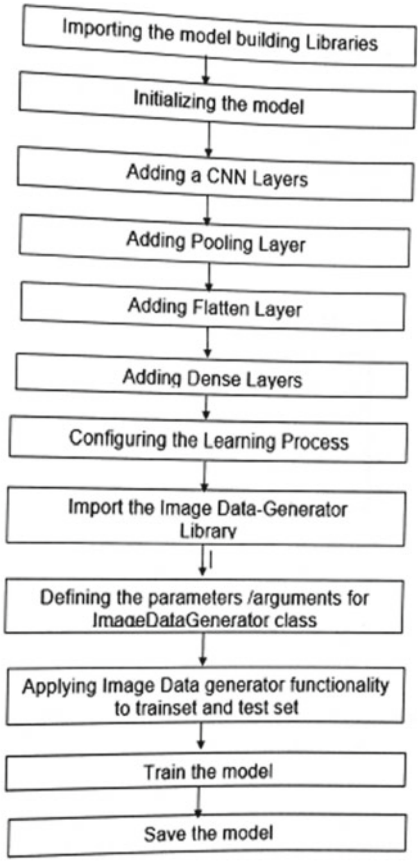


Fig. 10 Flow chart of the algorithm

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

Fig. 11 Confusion matrix

Fig. 12 Accuracy

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

Fig. 13 Precision

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

(True). False is because the model has anticipated incorrectly and positive because the class anticipated was a positive one.

False Negatives (FNs): FNs are the instances whilst the true magnificence of the records factor turned into digit one (True) and the anticipated is likewise is digit zero (False). False is due to the fact the version has anticipated incorrectly and poor due to the fact the elegance anticipated become a poor one.

Accuracy: Accuracy is category troubles, is the quantity of accurate conjectures made via way of means of the version over all forms of predictions made (Fig. 12).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

Accuracy: Accuracy is a superb degree whilst the goal variable lessons with inside the records are almost balanced. The version with inside the task is checked in opposition to the accuracy degree and we are reaching accuracy of 80%.

Precision: Precision is the percentage of efficaciously expected effective findings to the full expected effective findings (Fig. 13).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

5.2 Outputs

- **Video Processing**

  See Fig. 14.

- **Image Capturing**

  See Fig. 15.

Fig. 14 OpenCV code to process the video

Fig. 15 Image captured by camera

- **Alerting the Police**

  See Figs. 16 and 17.

6 Conclusions and Future Enchancements

6.1 Conclusion

Being secure and stable is the call for of the day. Our attempt in the back of this challenge is to layout and fabricate an answer that is so compact in itself that offer benefit of the safety device. This layout will cope with maximum of the essential problems confronted through ladies and could assist them to be stable. Existing structures offer the mechanism to music a man or woman with a safety device. The proposed mechanism presents viewing the region of the sufferer in phrases of range and longitude that may in addition be tracked the usage of Google maps with none safety device. This device enables to lower the crime charge towards ladies. Women's safety is an essential trouble in modern situation. These incidents may be delivered to a stop with the assist of actual time execution of our proposed device.

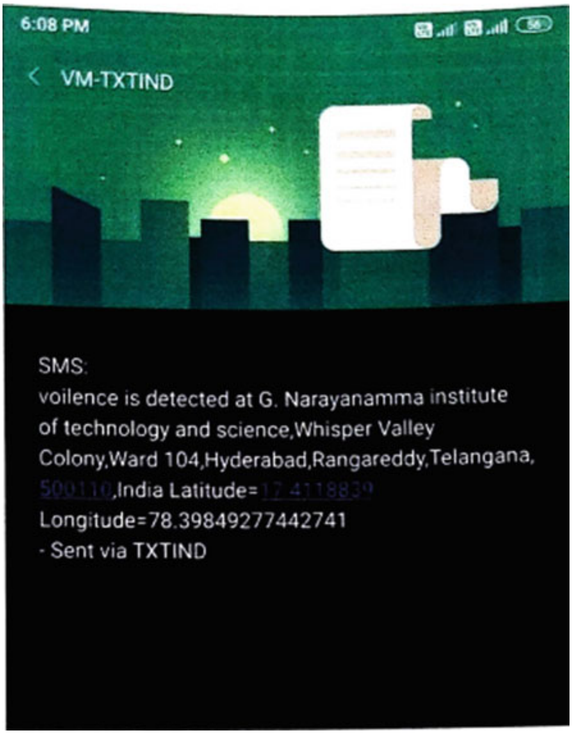


Fig. 16 Alert message sent through SMS

Fig. 17 Alert message sent through E-mail

6.2 Future Enchancements

As the technological modifications or new requirement from person to decorate the capability of product might also additionally calls for new edition to introduce. Although, the system is whole and running efficiently, new modules which decorate the gadget capability may be introduced with none important modifications to the complete gadget. Hence, the improvement era makes the gadget greater sturdy and reliable. As the brand new modules offer the capability which decorate the protection and security. Thus, it allows to meet the motive of the project.

Intelligent Alert Alarm System: We can further enhance the proposed model by including an alert system to it using Web browser library to create a Python alarm which would alert people around the place which will help for further fast help.

Primary School Children Safety: The college kids protection is predominant worries for mother and father in addition to college control because of the current incidents of toddler crimes like kids missing, abuse, etc. This module video display units the kid protection. Once they any form of combating among kids or any crook pastime in opposition to them it might ship the well-known message is ship to the major without delay which could offer an offer protection environment.

References

1. https://ieeeprojectsmadurai.com/IEEE%202019%20IOT%20BASEPAPERS/2_WOMEN%20SAFETY.pdf
2. https://www.ripublication.com/irph/ijict_spl/ijictv4n8spl_01.pdf
3. https://www.irjet.net/archives/V8/i5/IRJET-V8I5100.pdf
4. https://www.ijert.org/research/iot-based-artificial-intelligence-women-protection-device-IJERTCONV6IS13144.pdf
5. https://medium.com/@prathamanchan22/send-text-message-sms-in-python-using-fast2sms-al399e863552
6. https://docs.fast2sms.com/#get-method
7. https://www.geeksforgeeks.org/how-to-get-geolocation-in-python/
8. https://docs.opencv.org/4.5.2/d0/de3/tutorial_py_intro.html
9. https://en.wikipedia.org/wiki/Convolutionalneuralnetwork
10. https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/
11. https://towardsdatascience.com/1Q-minutes-to-building-a-cnn-binary-image-classifier-in-tensorflow-4e216b2034aa?gi=b93cb2f7del8
12. https://www.w3schools.com/python/refstringformat.asp
13. https://pythonprogramming.net/loading-images-python-opencv-tutorial/
14. http://www.ir.juit.ac.in:8080/jspui/bitstream/123456789/15835/1/121095.pdf

Emotion Detection on Twitter Text Using Machine Learning Techniques with Data Augmentation

Hemantha Kumar Kalluri, Karthik Kotam, Harish Thota, Roja Kuchipudi, Sanjana Sai, and P. Krishna Prasad

Abstract Social webs like Instagram, Twitter, and WhatsApp are full of deliberations involving sentiments, feelings, and impressions of human beings worldwide. Moreover, understanding and segregating texts based on emotions is a complex task that could be considered progressive sentiment analysis. As sentiments play a crucial role in human interaction, the skills to perceive it through textual content analysis has numerous applications in natural language processing (NLP) and human–computer interaction (HCI). This paper suggests classifying and examining tweets based on six basic emotions: happiness, fear, anger, disgust, surprise, and sadness. Language translators are used to apply data augmentation. Experimental results show that augmented data provides better results than the original data.

Keywords Emotion detection · Classification · Analysis

H. K. Kalluri (✉) · K. Kotam · H. Thota · R. Kuchipudi · S. Sai · P. Krishna Prasad
CSE Department, SRM University AP, Guntur, Andhra Pradesh, India
e-mail: hemanthakumar.k@srmap.edu.in

K. Kotam
e-mail: kotam_karthik@srmap.edu.in

H. Thota
e-mail: thota_harish@srmap.edu.in

R. Kuchipudi
e-mail: kuchipudi_roja@srmap.edu.in

S. Sai
e-mail: sai_sanjana@srmap.edu.in

P. Krishna Prasad
e-mail: krishnaprasad.p@srmap.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_76

1 Introduction

Emotions are detailed as subtle to extreme feelings about something resulting from internal or external events that have a compelling encounter with a person. In today's world, social media has become an important channel for people to convey their thoughts and opinions. Millions of humans avail social media to share, upload and comment on events or news stories across the globe. Tweets represent people's counter to different opinions, circumstances, and situations. Business analysts can utilize this data to record people's feelings and thoughts about their materials and goods. Vendors, for instance, use social media websites such as YouTube, Instagram, and Twitter for advertising product data and productively gathering client feedback in the business world [1]. People's feedback is essential for business marketers to calculate client satisfaction and keep track of the competitors, as well as for clients who want to study more about a good or assistance beforehand buying it. Human language understanding and creation are two components of natural language processing (NLP). NLP comprises two essential areas: sentiment analysis and emotion recognition. Sentiment analysis discloses whether the data is positive, neutral, or negative. On the other hand, emotion detection is a means of identifying different human emotions, such as happiness, anger, or sadness.

Most studies on sentiment analysis [1] tells only whether people's reactions are positive or negative. It does not give supplementary information like specifying the client's definite feelings or the degree of their response. Emotion detection does not ponder whether the selected emotions are good or bad but rather targets analyzing the exact human emotion present in a text, image, video, or voice recording. Humans can decipher the sentiments of those around them through numerous signs, including body language, voice tone, facial expressions, and words exchanged.

2 Literature Study

Krommyda et al. [1] proposed a rule-based emotion classification algorithm. Plutchik et al. [2] published a book on theory, research, and experience on Emotion. Gaind et al. [3] proposed a method to classify text into six different emotion categories: happiness, sadness, fear, anger, surprise, and disgust. The researchers used two different approaches and combined them to extract these emotions from the text effectively. The first approach is based on natural language processing and uses several textual features like emoticons, degree words and negations, parts of speech, and other grammatical analyses. The second approach is based on machine learning classification algorithms. The researchers conducted experiments for classifying tweets taken from Twitter.

Having read recent research papers [1, 4–10] and their analysis gave us some methods of approach using both machine learning and deep learning in emotion detection. They created three massive datasets labeled with the emotions conveyed via hashtags using the convolution neural networks (CNN) algorithm and proposed a unifying model for recognizing multiple emotions. A total of eight emotions have been identified. Using the updated hashtag emotion database, performance has improved, but there was no semantic analysis performed, and just static data was used for analysis. For labeling, only hashtags were utilized. Later, emotion-word hashtags were operated using the support vector machine (SVM) algorithm to label tweets manually.

3 Proposed Method

Special characters (e.g., #, ?, !, etc.) and emoticons are frequently found in the raw data acquired from Twitter. As a result, those special characters have been removed.

1. Tokenization: It is breaking down a text into separate words. Tokenization is the term for this procedure. These procedures are completed to clean the data and prepare it for analysis.
2. Remove stopwords: Stopwords are words that appear frequently but have no bearing on the content of a page (e.g., a, an, the, etc.). As a result, stopwords are removed.
3. Lemmatization: It refers to performing things accurately using both morphological analyses of words and vocabulary to eliminate only articulation endings and return the word in the base form only. It is called the lemma.

3.1 Dataset Preprocessing

The machine learning methodology is one method for sentiment analysis. The whole dataset is separated into two parts: one is the testing dataset, and the other is the testing dataset. The training dataset is used to train the model by providing the properties of several occurrences of an item. After that, the testing dataset is utilized to regulate how well the model was trained that is obtained from the training dataset.

The naive Bayes method is a supervised algorithm for pointing to classification concerns solely based on the Bayes theorem. Bayes' theorem determines the likelihood of an event recurring, where the probability of an earlier event is given. The below mathematical equation discloses Bayes' theorem:

$$P(h|d) = (P(d|h)P(h))/P(d) \tag{1}$$

where

$P(h|d)$ is the posterior probability of the class.
$P(h)$ is the prior probability of class.
$P(d|h)$ is the likelihood which is the probability of the predictor given class.

$P(d)$ is the predictor's prior probability.

Decision trees are a supervised nonparametric learning method that can be used for all classification and regression works. Classification trees are a tree layout in which the objective variable can return a distinct set of values. The K-nearest neighbors (KNN) algorithm is used to find the query object belonging to which class. The mathematical formula for finding the Euclidean distance is given in Eq. 2.

$$d = \sqrt{\left((x2 - x1)^2 + (y2 - y1)^2\right)} \tag{2}$$

XGBoost is a classification algorithm that performs optimally distributed gradient-boosting machine learning methods using the gradient boosting framework. XGBoost is a one-part method of ensemble learning. Ensemble learning is an approach for methodically linking the predictive capabilities of most learners. The result, in the end, is a single model that amalgamates the outputs of various models.

Logistic regression is an approach used to foresee a dependent variable when a set of independent variables are given such that all the dependent variables can be categorized. It is utilized in statistical software to comprehend the association between the dependent variable and the independent variables by evaluating tendencies using an equation of logistic regression.

Bag of words: It is a feature extraction technique that converts text input into numerical vectors that can be used as features. Each word (token) in a document is counted using those numbers. Produce a sparse matrix (mainly 0s). It is creating the corpus of sentences and converting them into lowercase to not differentiate between the exact words. It is creating the corpus of sentences and converting them into lowercase to differentiate the exact words. Developing a vocabulary requires identifying unique words, setting an index for the vocab words, and creating a sparse matrix.

The count vectorizer library from Scikit-learn must be fitted and modified on the corpus. CountVectorizer(), as below, provides specific arguments which enable the performance of data preprocessing, such as stop_words, token_pattern, and lower. The token size is adjustable, ngram_range says one word, but this can be changed depending on the use case. The token size is adjustable, as the default ngram_range says one word, but this can be changed depending on the use case.

A mathematical metric called frequency-inverse document frequency is stated to designate how important a word is to a folder, file, or document in a corpus or a collection. Since it appoints a numeric to a term based on its relevance in a file scaled by the priority of all texts in your group, TF-IDF is the most used approach for weighing terms for NLP tasks. Generally occurring texts in the English language, choose words that are most descriptive of your text. The frequency with which a term, t, is measured by TF:

$$[[tf]]\_(t, d) = n\_(t, d)/(\text{no. of terms in document}) \tag{3}$$

$n_{t,d}$ is how frequently the term "t" appeared in the file/document "d" in the numerator. So, each document and word will be prescribed an exclusive TF value. The IDF is a metric for understanding the effectiveness of a phrase. The IDF value is so required because only computing the TF is not at all sufficient to understand the significance of words:

$$[[idf]]\_t = \log((\text{no, of documents})/(\text{no. of documents with term '}t^{'})) \quad (4)$$

We can now compute the TF-IDF score for each word in the corpus. Higher-scoring words are more important, while lower-scoring words are less important:

$$[[(tf\_idf)]]\_(t, d) = [[tf]]\_(t, d) * [[idf]]\_t \quad (5)$$

The TF-IDF scores for our vocabulary have now been acquired.

3.2 Data Augmentation

To artificially enhance the dimensions of the actual dataset, we use data augmentation techniques to generate different versions. Data augmentation is used in natural language processing (NLP) models to deal with a lack of data diversity and paucity of data.

Random under-sampling and random over-sampling are two primary strategies for dealing with an issue of class imbalance. The above-mentioned problem can be solved intuitively by summing samples to minority classes, eliminating samples from majority classes, or mixing the two. The majority class is randomly sampled. This is known as random under-sampling. This step is recursively done until harmony is acquired in terms of the minority class. While this procedure aids in the harmony of both the minority and majority classes, important information is likely lost when samples from the majority class are removed. The term "random over-sampling" refers to duplicating samples from the minority group. The additional step is recursively done until the elements in the majority class are balanced; nonetheless, the program may cause the model to be over-fitted to the minority class.

Using Google Translate as a tool: We can see if training the model in a single language is better or worse than training in numerous languages by experimenting. We can translate sentences randomly into another language and then back to the original. We will translate into languages already in our dataset, although translating into languages that are not in our dataset could improve our performance.

We have implemented all the algorithms using a "bag of words" (count vectorization). Later taken the same datasets and applied TD-IDF and got less accuracy as it ignored the essential words. So, to get higher accuracy, use data argumentation and make a new dataset by combining the initial datasets and by Google translation sentences. The dataset1 (D1) before data argumentation consists of 39,241 tuples; after augmentation, the dataset holds 90,719 tuples, and it is assumed as dataset2 (D2) (Table 1).

Table 1 Number of tuples before and after data augmentation

Sentiment	Count actual	Count after augmentation
Neutral	8638	8638
Worry	8459	8459
Sadness	5209	5209
Love	5165	7677
Surprise	3842	8436
Fun	2187	8195
Relief	1776	7630
Hate	1526	8887
Empty	1323	5789
Enthusiasm	827	6072
Boredom	179	6980
Anger	110	8747
Total	39,241	90,719

4 Results

After implementing the proposed algorithms and data argumentation, we have obtained the following results. The classification algorithms were applied to all emotions and only to three emotions (positive, negative, and neutral) on dataset1 (D1) and dataset2 (D2). D1 indicates the dataset1, and D1(3) indicates the positive, negative, and neutral emotions in dataset1. The experimental results are given in Table 2.

The experimental results show that the accuracy is improved with the proposed data augmentation in all the classification algorithms.

Table 2 Experimental results of before and after augmented data

Classification algorithm	Accuracy on original dataset		Accuracy after data augmentation	
	D1	D1(3)	D2	D2(3)
Naïve Bayes	31	72	52.05	80.31
Logistic Regression	33	88	58.67	92
KNN	31	72	48.46	67
Decision tree	27	87	55.29	90.5
XGBoost	34	85	32.9	73.6

5 Conclusion

Emotion recognition is a large field of study that has found applications in critical sectors such as diagnosing psychological problems in people, such as anxiety or depression, evaluating the public mood of a community, and sentiment analysis. The suggested technique is validated using real-time tweets acquired from Twitter. This model outperforms the prior unison model by a wide margin and completes the task in a significantly shorter time. The proposed method can be used for a variety of tasks, including automatic real-time online user feedback analysis, user reputation monitoring, etc.

References

1. Krommyda M, Rigos A, Bouklas K, Amditis A (2020) Emotion detection in Twitter posts: a rule-based algorithm for annotated data acquisition. In: 2020 International conference on computational science and computational intelligence (CSCI), pp 257–262.https://doi.org/10.1109/CSCI51800.2020.00050
2. Plutchik R (1980) Emotion: theory, research, and experience. Academic Press
3. Gaind B, Syal V, Padgalwar S (2019) Emotion Detection and analysis on social media. arXiv:1901.08458v2 [cs.SI] 12 Jun 2019
4. Deshpande A, Paswan R (2020) Real-time emotion recognition of Twitter posts using a hybrid approach. ICTACT J Soft Comput 10(4):2125–2133
5. Polignano M, de Gemmis M, Semeraro G (2019) Emotion detection in conversations through Tweets, CNN and LSTM deep neural networks. In: Proceedings of the 13th international workshop on semantic evaluation
6. Shah FM et al (2019) Emotion detection from tweets using AIT-2018 dataset. In: 2019 5th International conference on advances in electrical engineering (ICAEE). IEEE
7. Krommyda M et al (2021) An experimental analysis of data annotation methodologies for emotion detection in short text posted on social media. Informatics 8(1)
8. Nandwani P, Verma R (2021) A review on sentiment analysis and emotion detection from text. Soc Netw Anal Min 11(1):1–19
9. Yousaf A et al (2020) Emotion recognition by textual tweets classification using voting classifier (LR-SGD). IEEE Access 9:6286–6295
10. Murthy AR, Anil Kumar KM (2021) A review of different approaches for detecting emotion from text. IOP Conf Ser: Mater Sci Eng 1110(1)

Leaf Disease Prediction and Detection Using CNN

R. Sravani, N. Sandhya, Kamuni Kavita, and Snigdha Soni

Abstract Agriculture is the prime sector of India's economy and stands as one among the top two crop producers in the whole world. Farm production is very important and plays a primary part in agribusiness sector as it contributes around 18.1% to the country's GDP. The less production of annual food rate is mainly dependent to infertile farmland and infected crops. The identification of crop infections in the area of agricultural sector is exceptionally demanding. When the detection is not accurate, the crop production may decrease and the market value of that product may affect a drastic loss. This indicates the need of new techniques which help in the detection of crop infections and its growth. Hence, we are proposing a new model by the use of convolution networks by classification of infected leaves. This proposed system can detect thirteen different and complex kinds of crop infections by differentiating them from healthy leaves. This proposed model of detection and prediction of plant diseases is mainly introduced to focus on the agricultural and agro-based business. This helps in the early detection and identification of crop diseases which helps the farmers to take necessary precautions or measures that in turn can help with the crop productivity.

Keywords Neural networks · SVM · Pooling layer · Image processing

R. Sravani · N. Sandhya · K. Kavita (✉) · S. Soni
BVRIT, Hyderabad, India
e-mail: Kavita4.Kamuni@gmail.com

R. Sravani
e-mail: sravani.r@bvrithyderabad.edu.in

N. Sandhya
e-mail: Sandhya.n@bvrithyderabad.edu.in

S. Soni
e-mail: Snigdhasoni@bvrithyderabad.edu.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_77

1 Introduction

The major problem of crop infections is closely linked with the problems of environmental and biological changes in the farmland which is not suitable for sustainable agriculture [1]. Much research has shown that the global warming and climate change will cause frequent changes in percentage of agent and antibody development which in turn cause infections. They have the capability to change the host resistance levels and end up with changes in physiological structure of host-pathogens [2, 3]. The current situations prove that a disease can transfer globally and can take various forms of structures. Different kind of infections are identified in the areas where there was no trace of diseases and there is no proper awareness to combat them [4–6]. Usage of pesticide without proper training and experience will result in the production of semipermanent protection against pathogens, which may in turn result in decreasing the ability to fight back against the infections. Correct, timely diagnosis and identification of crop disease are among the main goals of agricultural sector [7]. In this dynamic surroundings, sustainable and fast disease detection is very vital. There exist many methods to identify the pathologies of plants. Few diseases do not give any symptoms, and for a few, the symptoms appear too late. Nevertheless, many infections create a kind of similarities in this field, and thus, the Oculus identification is the main method adopted by many for disease identification. In order to achieve and identify the correct condition of plant, many plant specialists possess keen observation [8]. Differences in the symptoms which are indicated by many infected plants might cause associate uneven diagnosing since amateur gardeners and hobbyists may have additional difficulties deciding it than an expert plant pathologist. An automatic system designed facilitates establishing plant wellness by the plant's look, and visual symptoms can be of nice help to amateurs within the agriculture method and additionally trained professionals as a verification system in disease medicine. Advances in laptop vision gift a chance to expand and enhance the follow of precise plant protection and extend the market of PC vision applications within the field of preciseness agriculture. Exploiting common digital image process techniques like color analysis and threshold [9] were used with the aim of detection and classification of plant diseases. Numerous different approaches are presently used for three detective work plant diseases, and most types are artificial neural networks (ANNs) [10] and support vector machines (SVMs). They are combined with totally different ways of image preprocessing in favor of higher feature extraction. In machine learning and scientific discipline, ANN is an associate information processing paradigm that was impressed by the approach of biological nervous systems. The brain is composed of an oversized variety of complex interconnected elements called neurons which work together in order to resolve specific problems. The man-made nervous system cell may help as a processing component with various inputs and one ultimate output. Though many artificial neurons may give several outputs, the only ones with a precise and singles outputs are considered. Their inputs may withstand any value between zero and one. Also, the nerve cell has weights for each input associated with an overall bias. In addition we observed, it should be first point where you can recognize the diseases and is no longer helpful [11].

2 Literature Review

Based on the work and the results obtained by Garrett and Dendy, there are many ways for the leaf processing method to be implemented like RNA double-stranded review, sampling of organic compounds, etc., [1]. Among the few, we have chosen to use the image processing method. Various types of measures are in use to implement the identification of plant diseases using computer vision. One of such methods is by extracting the features of color. Based on the inputs from Coakley and Scherm, the noise from color extraction like using of camera flash and others were successfully identified and removed [2–4]. Bodhe and Patil have applied these detection methods of identification by applying it in sugarcane leaves and were able to achieve an accuracy of around 98.6%. Furthermore, this feature and texture extraction model was also implemented by Kumar and Patil where characteristics like homogeneity, inertia, and the similarities are obtained by implementing the gray-level matrix calculation co-occurring in the pictures of leaves [5]. They have performed many experiments on the maize leaves using color extractions. We get a stronger result in image enhancement by combining these previously mentioned methods and helps in more efficient classification [6, 7]. The authors even presented with a survey of many well-known ways for extracting the features. These methods were implemented along with the advanced development in the field of AI, and it is research based on various methodologies. There were various methods for including the feature extraction and the neural network ensemble (NNE) in crop disease detection. Training by the use of neural networks and using their results for training and analysis provides a better result for prediction. This method was very efficient for only the herbal and biological infections with an accuracy of around 91% [8]. We propose a system that is purely based on the deep learning approach and its application. This takes the input of various healthy and infected leaves data as images for building the model using CNN, and when a test image 5 is passed for detection, the test image is processed based on feature extraction and prediction of infection is done to give appropriate results. This provides an efficient approach based on proven image processing techniques with better accuracy for identifying crop diseases using image processing [10].

2.1 Proposed Method

A computer-based model can be used to maintain agricultural health and encourage a healthy cultivation where different kinds of infectious symptoms can be identified. A tool can be used by both professionals and amateurs in farming. A system which

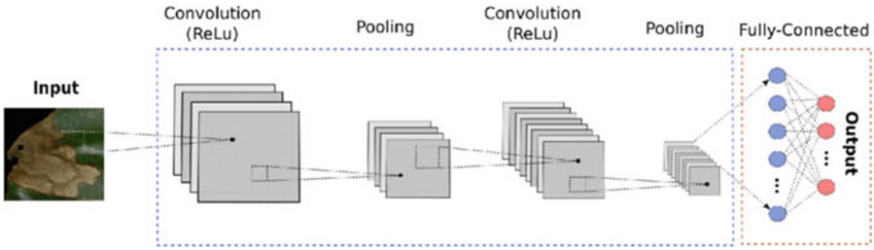


Fig. 1 Block diagram for CNN

can help farmers in taking necessary precautions and protect their crop throughout the year when the crop is ready for cultivation. Advancement in the field of computer vision and artificial intelligence has provided a way for many to design tools which can be extremely useful in many fields in our everyday life. This provides a plan to implementing accurate planting and monitoring the crop in order to achieving precision farming. Threshold and color analyses [9] are most popular image processing methods to detect and classify diseases in leaves. After that, a variety of methods came into existence based on artificial neural networks (ANNs) and support vector machines (SVMs) [10]. Here, a minute change of image preprocessing is implemented to extract higher level features. These artificial neural networks work similar to a human brain where highly functional neurons are interconnected which help in information processing. These neurons work with multiple inputs and a single output. So, every input comprises of a certain bias.

Classification of layers in CNN: A convolution neural network comprises of 1 input layer, 2 convolution function layer, 3 activation function layer, 4 fully connected layer, and 5 pool layer (Fig. 1).

2.2 Dataset Description

The dataset used for processing comprises of about 87,000 RBG pictures of diseased and healthy crop leaves which are categorized into 38 different set of classes. The whole dataset is segmented into 80/20 ratio of training and validation set maintaining the directory structure. For the prediction purpose, a new directory consisting of 33 test images is created. The dataset covers a variety of plant diseases including apple scab, black rot, cedar apple rust, cherry mildew, corn Cercospora leaf spot, corn gray leaf spot, corn common rust, grape black rot, grape esca (Black measles), grape Isariopsis leaf spot, orange citrus greening, peach, tomato, and bell pepper bacterial spot, potato early bright and late bright, strawberry leaf scorch, tomato early bright and late bright, tomato leaf mold, Septoria leaf spot, spider mites, target spot, mosaic virus, yellow leaf curl virus, etc.

3 Implementation

3.1 Problem Definition

A system is designed to detect and identify the infected crop. This proposed system can detect 13 different and complex kinds of crop infections by differentiating them from healthy leaves. This proposed model of detection and prediction of plant diseases is mainly introduced to focus on the agricultural and agro-based business. This helps in the early detection and identification of crop diseases which helps the farmers to take necessary precautions or measures that in turn can help with the crop productivity.

3.2 Detection with Image

The model is trained using a dataset with about 87K RBG images of healthy and infected leaves which are divided into 38 different set of classes. So once the training is done, we can use this model for prediction. The test set for prediction consists of about 33 different varieties of leaves, both healthy and unhealthy. When a test image of infected leaf is passed, prediction.py module is used in the backend for detection and returns the most accurate result using convolution neural networks.

3.3 Detection with Camera

This model allows real-time capturing of images using a camera. In this case, the inbuilt camera of the working laptop is considered to capture the image. When an image of infected leaf is presented as an input to the system, the image is used to prediction task. This is used in the Cnn_predict.py module. This is considered as a test image and is compared with the training set that is used to building the model. The prediction works similar to the detection with image module using prediction.py in the backend. After processing, the appropriate results are returned on the screen.

3.4 Build CNN Model

The user interface consists of build CNN model button which provides the functionality of building a convolution neural network model using the data that is fed to the system. The module CNN.py defines the constitutional base using a very common pattern, i.e., a stack of Conv2D and MaxPooling2D layers. As an input, convolution neural network takes tensors of shape (image height, image width, and color channels) while ignoring the batch size. Color_channels refer to (R, B, G). The output is a three-dimensional tensor of the shape height, width, and channels. In order to finish the building of model, we feed the tensor which we got in as previous output from the convolution base into denser layers to perform the task of classification. These denser layers will take vectors as an input which are usually in one dimension while the existing output is a three-dimensional tensor. So, we flatten 3D output to 1D using models add(layers.Flatten()) and add few more dense layers on top of the built tensor.

4 Results and Discussion

This chapter presents results of experiments including the build of model and prediction of leaf disease. As a result of the proposed model, figure illustrates the admin home page which displays the built and detection modules that are present in the application (Fig. 2).

When the user clicks on the detection with image button, the application will redirect the user to the detection page as shown in Fig. 3 where the user can load the image of the infected plant leaf that the user wishes to detect. After inserting the image from local drive, the name of the file is shown on the screen along with the path of the file.

Once the user clicks on detect button, the system will process the test image and run the CNN module to detect the type of disease that is infected by the plant. Once the detection is completed, the result is displayed on the screen (Figs. 4 and 5).

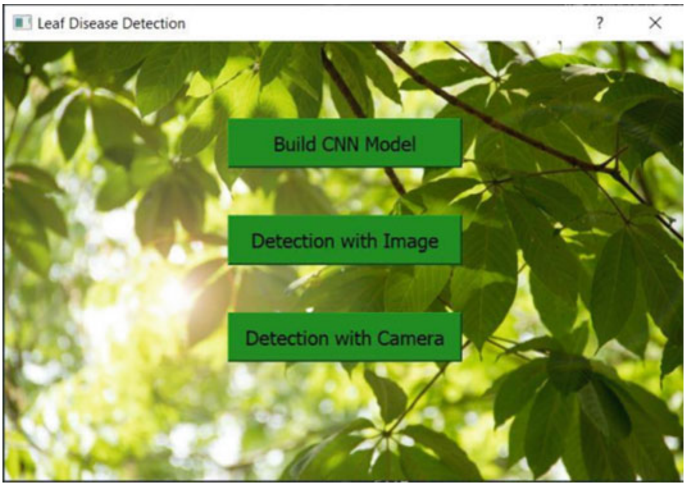


Fig. 2 Home page displaying the built and detection modules

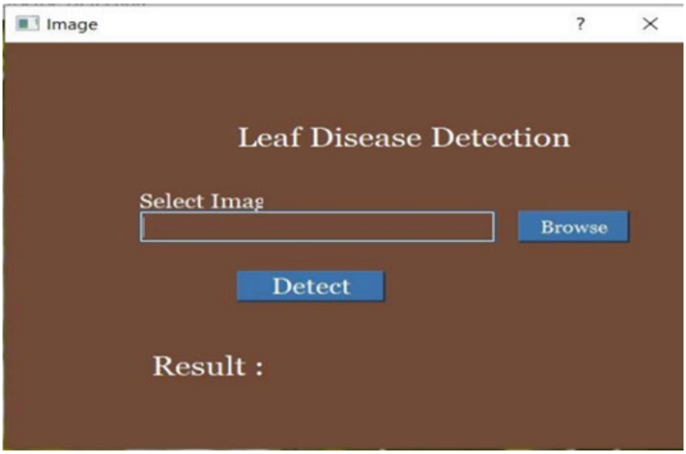


Fig. 3 Detection using image module where the user can insert the image of infected plant leaf for disease detection

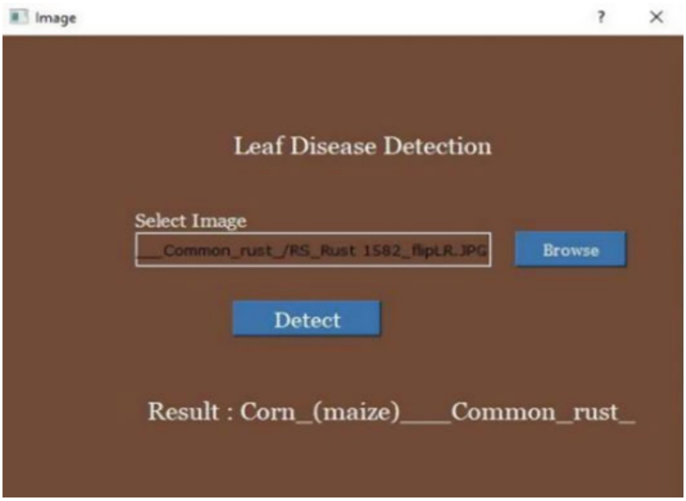


Fig. 4 Screen displaying the result of detection including the type of leaf and the disease it is infected with user interface for disease detection

```
Corn_(maize)___Common_rust_
1. Corn_(maize)___Common_rust_: 99.999511%
2. Apple___Apple_scab: 0.000317%
3. Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot: 0.000163%
4. Potato___Early_blight: 0.000005%
5. Apple___Cedar_apple_rust: 0.000003%
res= Corn_(maize)___Common_rust_
```

Fig. 5 Accuracy of prediction

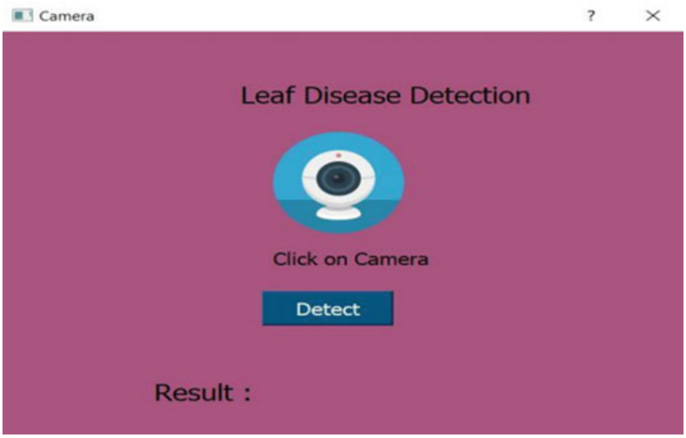


Fig. 6 Detection using camera module where the system captures the image used for prediction of plant disease

We can also perform the prediction without uploading an image into the system by using the camera-based detection as shown in Fig. 6. Using the camera, the system captures the image and uses it as a test data for the prediction task.

5 Conclusion and Future Work

5.1 Conclusion

There are various approaches introduced in detection and analysis of machine based or computer vision-based techniques for plant disease, but there is still a lack in the part of efficiency for a few. An idealistic approach for this detection using deep learning is introduced here which helps in recognizing crop disease using the image of an infected leaf. The model developed is also capable to identify the existence of a virus and differentiate between infected leaves and healthy ones. The whole procedure was from the grouping of the pictures which are used for the process of training and testing for preprocessing followed by augmentation of the image, finally performing deep CNN and fine-tuning procedure of jobs. In order to achieve greater efficiency, many different tests were performed using various test sets of data including various images of infected and healthy plant leaves. This proposed model's final accuracy is estimated to be 96.3%. Hence, we are able to successfully identify the plant diseases which can be very useful in agricultural sector to identify the infected crop and take necessary precautions or measures in order to increate crop productivity.

5.2 Future Work

As a future enhancement to this current model, many varieties of plant and their associated diseases can be added by collecting data and including them for detection and prediction. A bigger data model gives us more scope to detect different kinds of plant infections. We can implement this image-based plant disease detection for precision farming by empowering it with robotics or drone to increase the crop productivity by capturing the images in real time used to analysis.

References

1. Garrett KA, Dendy SP, Frank EE, Rouse MN, Travers SE (2006) Climate change effects on plant disease: genomes to ecosystems. Annu Rev Phytopathol 44:489–509
2. Coakley SM, Scherm H, Chakraborty S (1999) Climate change and plant disease management. Annu Rev Phytopathol 37(1):399–426
3. Chakraborty S, Tiedemann AV, Teng PS (2000) Climate change: potential impact on plant diseases. Environ Pollut 108(3):317–326
4. Tatem AJ, Rogers DJ, Hay SI (2006) Global transport networks and infectious disease spread. Adv Parasitol 62:293–343
5. Rohr JR, Raffel TR, Romansic JM, McCallum H, Hudson PJ (2008) Evaluating the links between climate, disease spread, and amphibian declines. Proc Natl Acad Sci USA 105(45):17436–17441
6. Van der Zwet T (2001) Present worldwide distribution of fire blight. In: Proceedings of the 9th international workshop on fire blight, vol 590, Oct 2001
7. Miller SA, Beed FD, Harmon CL (2009) Plant disease diagnostic capabilities and networks. Annu Rev Phytopathol 47:15–38
8. Riley MB, Williamson MR, Maloy O (2002) Plant disease diagnosis. The Plant Health Instructor
9. Arnal Barbedo JG (2013) Digital image processing techniques for detecting, quantifying and classifying plant diseases, vol 2, article 660. SpringerPlus, pp 1–12
10. Cartwright H (2015) Artificial neural networks. Humana Press
11. Vijayendar Reddy Ch, Obulakonda Reddy R, Srinivasulu B (2021) Image based plant disease detection using ML algorithm, Issue 8, pp 10758–10767

Implementation of a Hybrid Movie Recommendation System for Large-Scale Data

T. Adilakshmi, S. Sree Lakshmi, T. Anuvik, and L. Rahul

Abstract Recommender systems are being used extensively in the modern digital age as mammoth of user data is available in the contemporary private databases. Handling large amounts of data has become stressful due to proliferation of user data available on the Web. The existing recommender systems are built using algorithms like K-means, collaborative filtering, etc. Recommendation on big datasets where content is not given needs to be replaced and work in parallel with tag-genome data. In this study, we used CURE algorithm for clustering the data. Pre-eminent clustering algorithms are either bent toward spherical shape clusters with close sizes or sensitive with respect to others. Here, we are introducing and implementing a new methodology which is more suitable to handle the outliers which even suggests clusters of non-spherical shape of heterogeneous sizes. For managing voluminous databases, CURE implements a mix of partitioning and random sampling.

Keywords Movie recommendation system · CURE · Parallelization

1 Introduction

Recommendation system helps in prediction of the rating of the item or to suggest item to the user.

Goals of Recommendation System

1. **User-Experience Customization**: Enhances the hands-on experience for the user by analyzing the user's behavior and preferences.
2. **Optimum Item Search**: Splits the items based on the parameters.

T. Adilakshmi · S. Sree Lakshmi (✉) · T. Anuvik · L. Rahul
Vasavi College of Engineering, Hyderabad 500031, India
e-mail: s.sreelakshmi@staff.vce.ac.in

A. Kumar et al. (eds.), *Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-99-2742-5_78

1.1 Content-Based and Collaborative-Based Recommendation System

In content-based methodology, the system suggests user's preferred item based on their selection criteria by using similarity. It suggests items based on their existing functionalities. It recognizes the closeness among the items on the basis of content. **Collaborative-based** methodology suggests items based on collaboration of users which are luminous methods used in market.

1.2 Hybrid Recommendation System

Hybrid recommendation system is an amalgamation of content- and collaborative-based recommendation systems. We used hybrid recommendation system in our project to recommend movies to user. Here, we used the dataset which comprises of 1128 tag relevance values (Movie belongingness to a particular tag) and 228 movies. Tag genome records indicate how strongly a tag is related to item ranging from 0 to 1, where '0' is the weakest and '1' is the strongest. Here, we used feature reduction for reducing 1128 tags to 300 clustered tags. The clustering process here is done using Pearson correlation metric using agglomerative clustering.

1.3 Parallelization

Spark is known for its parallel processing, which means a data frame or a resilient distributed dataset (RDD) is being distributed across the worker nodes to gain maximum performance while processing. Using Databricks, we implement multi-threading which is relatively quick to set up compared with other optimization methods. The improvement could be unlimited if we have a large enough cluster and plenty of jobs to run parallelly. The purpose of using multi-threading is not only to save time, but also to fully utilize the clusters' compute power to save cost by finishing the same amount of jobs within less time, or within the same amount of time on a smaller cluster, which gives us more options to manage the end-to-end pipeline.

2 Recommendation Using Clustering

Method of scalable collaborative filtering is proposed in [1]. Various recommendation methods are applied by clustering the data efficiently. The algorithms proposed in [2–4] efficiently discover clusters in large spatial database with noise. The approach proposed in [5] used SVD for clustering. The paper applies the approach presented in [4] to the large-scale data. The data used is a part of the data available at [6]. The approach is implemented and tested using Databricks.

2.1 CURE

The clustering algorithm used is clustering using representatives (CURE) [4]. The pivotal reason for opting this algorithm is its compatibility with Big Data. The Crux of this algorithm is to represent a particular cluster in terms for few representative points, thereby scaling the large dataset into few representative points which helps in handing gigantic amount of data [7]. Also CURE [4] incorporates a compression factor that reduces the cluster size to some imaginary level and detects anomalies, thereby providing effective clusters. This algorithm can be parallelized using distributed environment and process the application for Big Data.

2.2 Data Pre-processing

The core dataset which is used is the Movie-Lens Tag genome dataset. The dataset we used consists of 1128 of tags with their corresponding movies tag relevance scores. Since the use of ratings might make the recommendation system obsolete, we used factual data to identify the category of the movie. These 1128 tags consist of the mixture of components cumulatively calculated to give a single tag ID. This tag ID indicates its categorical entity for a corresponding movie. The relevance scores lie between the range of 0 and 1 (0 least relevant, 1 most relevant).

2.3 Feature Reduction

The starting step is the feature reduction, we reduced the 1128 tag features to few sets of clustered tags. The clustering process here is done using *Pearson correlation* metric using *agglomerative clustering* (Fig. 1).

2.3.1 Experimentation with 300 Clustered Tags

The features were further reduced to 300 clustered tags. A total of 1128 tags is reduced into 300 clustered tags. For 300 tags and 228 movies, we applied CURE [4] to cluster the movies.

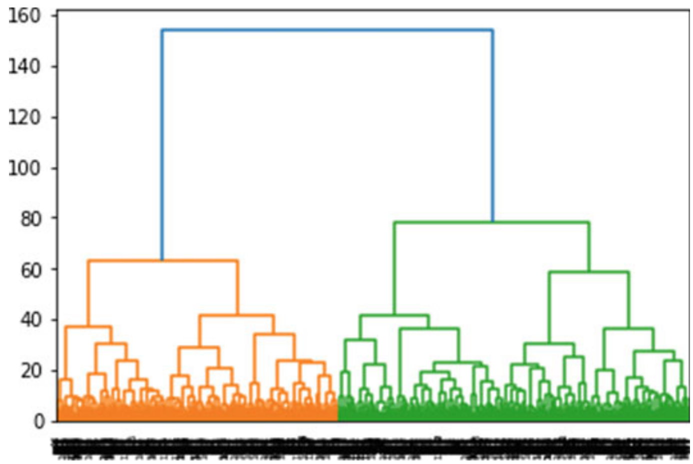


Fig. 1 Dendrogram showing the relationship between 1128 tags (tag relevance scores)

CURE Parameters

- Representative points = 5
- Alpha = 0.2
- Optimal number of clusters = 8

Elbow Point Using WSS Score

See Fig. 2.

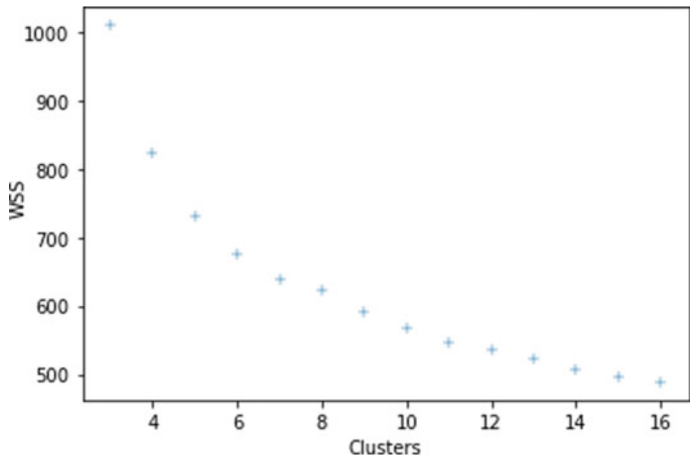


Fig. 2 Relation between within sum of clusters scores (WSS) on *X*-axis and number of clusters on *Y*-axis for 300 clustered tags

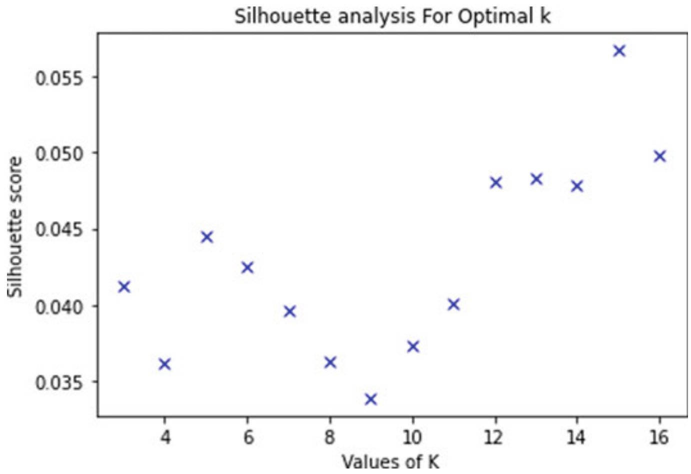


Fig. 3 Silhouette score analysis for 300 clustered tags, graph showing the silhouette scores on Y-axis and values of **K**

Silhouette Score Analysis

See Fig. 3.

(Number of clusters) on X-axis for 300 clustered tags.

The silhouette scores of the movie clusters for feature reduction to 300 clusters were positive, and hence, we reduced the tag scores to 300 clustered tags. From the WSS scores' graph plot, we obtained the optimal number of clusters, and the silhouette scores obtained from the 300 clustered tags are positive as shown in Fig. 9. We concluded that an optimal number of 300 clustered tags should be considered for further experimentation.

3 Proposed System

3.1 Scaling up the Process to 1000 User's Dataset

3.1.1 Databricks

Databricks is a centralized and open-source platform to process all the data. It provides an environment for interaction and scheduling of workloads of data analysis for data analysts, data engineers, and data scientists.

3.1.2 Configurations of Databricks

The Databricks runtime version, 10.4 LTS, includes Apache Spark 3.2.1 and Scala 2.12, driver type–15.3 GB memory, 2 cores, 1 Databrick unit (DBU).

3.1.3 Spark Data Frame

Firstly, to ensure parallelization, the data frame of the movie's dataset needs to be converted into spark data frame. The spark data frame internally divides it into multiple partitions, where different node are mapped to partition for handling. In this way, the parallelization is achieved, parts of the data frame is computed parallelly, and the results are merged to give the final output.

3.2 Algorithm

See Figs. 4 and 5.

Fig. 4 Working of the cure clustering algorithm [8]

```
procedure cluster(S, k)
begin
   T := build_kd_tree(S)
   Q := build_heap(S)
   while size(Q) > k do {
      u := extract_min(Q)
      v := u.closest
      delete(Q, v)
      w := merge(u, v)
      delete_rep(T, u); delete_rep(T, v); insert_rep(T, w)
      w.closest := x /* x is an arbitrary cluster in Q */
      for each x ∈ Q do {
         if dist(w, x) < dist(w, w.closest)
            w.closest := x
         if x.closest is either u or v {
            if dist(x, x.closest) < dist(x, w)
               x.closest := closest_cluster(T, x, dist(x, w))
            else
               x.closest := w
            relocate(Q, x)
         }
         else if dist(x, x.closest) > dist(x, w) {
            x.closest := w
            relocate(Q, x)
         }
      }
      insert(Q, w)
   }
end
```

Fig. 5 Merging phase of the algorithm

```
procedure merge(u, v)
begin
w := u ∪ v
w.mean := (|u|u.mean + |v|v.mean) / (|u| + |v|)
tmpSet := ∅
for i := 1 to c do {
    maxDist := 0
    foreach point p in cluster w do {
        if i = 1
            minDist := dist(p, w.mean)
        else
            minDist := min{dist(p, q) : q ∈ tmpSet}
        if (minDist ≥ maxDist){
            maxDist := minDist
            maxPoint := p
        }
    }
    tmpSet := tmpSet ∪ {maxPoint}
}
foreach point p in tmpSet do
    w.rep := w.rep ∪ {p + α*(w.mean-p) }
return w
end
```

3.3 Workflow of Hybrid Recommendation System

See Fig. 6.

3.4 Steps Followed in Recommendation System

Step 1: Tags cluster—Clustering the tag relevance values (Feature reduction) (Fig. 7).

Step 2: Implementing CURE [4] algorithm for clustering movies and user's dataset (Fig. 8).

Step 3: Building a recommendation engine to find the dominant genre for user's and movies clusters.

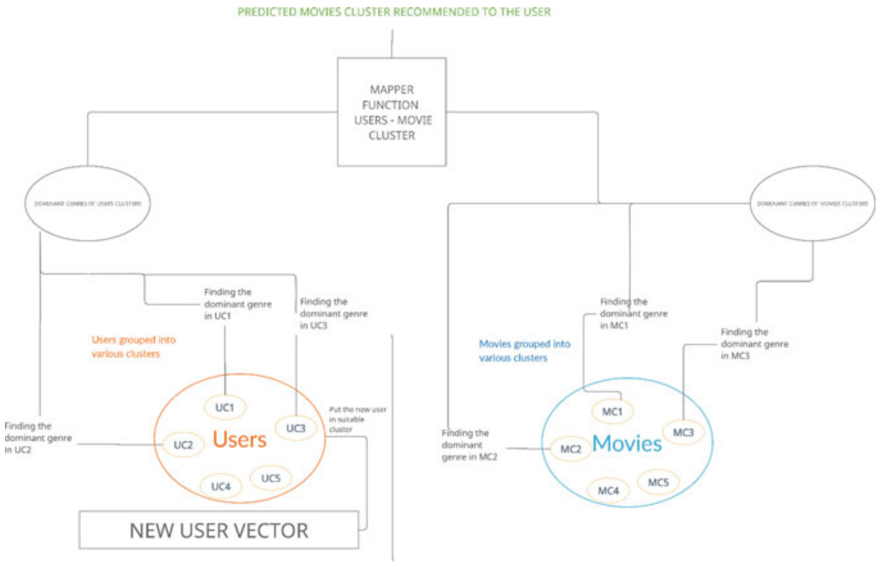
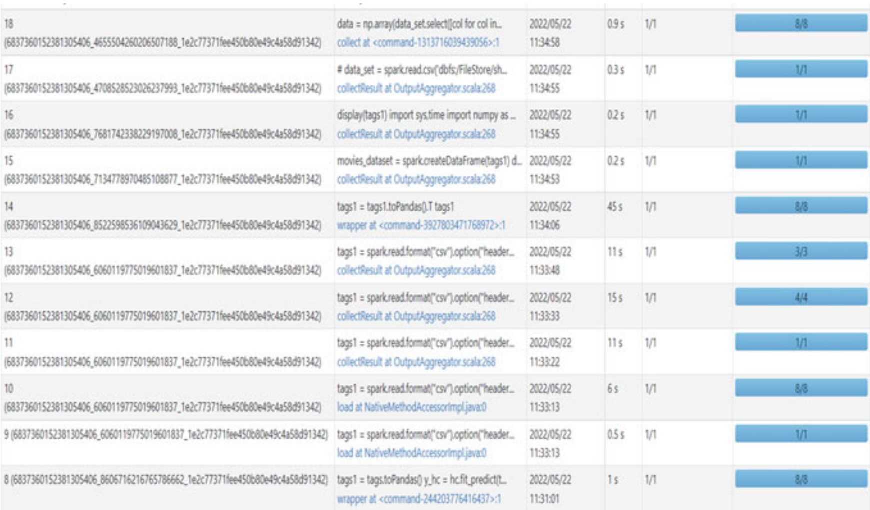


Fig. 6 Workflow of hybrid recommendation system

Fig. 7 Representing the formation tag clusters

```python
import numpy as np
import scipy.spatial.distance as distance
import sys

# Returns the distance between two vectors
def dist(vecA, vecB):
    return np.sqrt(np.power(vecA - vecB, 2).sum())

# This class describes the data structure and method of operation for CURE clustering.
class CureCluster:
    def __init__(self, id__, center__):
        self.points = center__
        self.repPoints = center__
        self.center = center__
        self.index = [id__]

    def __repr__(self):
        return "Cluster " + " Size: " + str(len(self.points))

    # Computes and stores the centroid of this cluster, based on its points
    def computeCentroid(self, clust):
        totalPoints_1 = len(self.index)
        totalPoints_2 = len(clust.index)
        self.center = (self.center*totalPoints_1 + clust.center*totalPoints_2) / (totalPoints_1 + totalPoints_2)

    # Computes and stores representative points for this cluster
    def generateRepPoints(self, numRepPoints, alpha):
        tempSet = None
        for i in range(1, numRepPoints+1):
            maxDist = 0
            maxPoint = None
            for p in range(0, len(self.index)):
                if i == 1:
                    minDist = dist(self.points[p,:], self.center)
                else:
                    X = np.vstack([tempSet, self.points[p, :]])
                    tmpDist = distance.pdist(X)
                    minDist = tmpDist.min()
                if minDist >= maxDist:
                    maxDist = minDist
                    maxDist = minDist
                    maxPoint = self.points[p,:]
            if tempSet is None:
                tempSet = maxPoint
            else:
                tempSet = np.vstack((tempSet, maxPoint))
        for j in range(len(tempSet)):
            if self.repPoints is None:
                self.repPoints = tempSet[j,:] + alpha * (self.center - tempSet[j,:])
            else:
                self.repPoints = np.vstack((self.repPoints, tempSet[j,:] + alpha * (self.center - tempSet[j,:])))

    # Computes and stores distance between this cluster and the other one.
    def distRep(self, clust):
        distRep = float('inf')
        for repA in self.repPoints:
            if type(clust.repPoints[0]) != list:
                repB = clust.repPoints
                distTemp = dist(repA, repB)
                if distTemp < distRep:
                    distRep = distTemp
            else:
                for repB in clust.repPoints:
                    distTemp = dist(repA, repB)
                    if distTemp < distRep:
                        distRep = distTemp
        return distRep

    # Merges this cluster with the given cluster, recomputing the centroid and the representative points.
    def mergeWithCluster(self, clust, numRepPoints, alpha):
        self.computeCentroid(clust)
        self.points = np.vstack((self.points, clust.points))
        self.index = np.append(self.index, clust.index)
        self.repPoints = None
        self.generateRepPoints(numRepPoints, alpha)
```

Fig. 8 Representing the implementation of CURE class



18 (6837360152381305406_4655504260206507188_1e2c77371fee450b80e49c4a58d91342)	data = np.array(data_set.select([col for col in... collect at <command-1313716039439056>:1	2022/05/22 11:34:58	0.9 s	1/1	8/8
17 (6837360152381305406_4708528523026237993_1e2c77371fee450b80e49c4a58d91342)	# data_set = spark.read.csv('dbfs:/FileStore/sh... collectResult at OutputAggregator.scala:268	2022/05/22 11:34:55	0.3 s	1/1	1/1
16 (6837360152381305406_7681742338229197008_1e2c77371fee450b80e49c4a58d91342)	display(tags1) import sys,time import numpy as ... collectResult at OutputAggregator.scala:268	2022/05/22 11:34:55	0.2 s	1/1	1/1
15 (6837360152381305406_7134778970485108877_1e2c77371fee450b80e49c4a58d91342)	movies_dataset = spark.createDataFrame(tags1) d... collectResult at OutputAggregator.scala:268	2022/05/22 11:34:53	0.2 s	1/1	1/1
14 (6837360152381305406_8522598536109043629_1e2c77371fee450b80e49c4a58d91342)	tags1 = tags1.toPandas().T tags1 wrapper at <command-3927803471768972>:1	2022/05/22 11:34:06	45 s	1/1	8/8
13 (6837360152381305406_6060119775019601837_1e2c77371fee450b80e49c4a58d91342)	tags1 = spark.read.format("csv").option("header... collectResult at OutputAggregator.scala:268	2022/05/22 11:33:48	11 s	1/1	3/3
12 (6837360152381305406_6060119775019601837_1e2c77371fee450b80e49c4a58d91342)	tags1 = spark.read.format("csv").option("header... collectResult at OutputAggregator.scala:268	2022/05/22 11:33:33	15 s	1/1	4/4
11 (6837360152381305406_6060119775019601837_1e2c77371fee450b80e49c4a58d91342)	tags1 = spark.read.format("csv").option("header... collectResult at OutputAggregator.scala:268	2022/05/22 11:33:22	11 s	1/1	1/1
10 (6837360152381305406_6060119775019601837_1e2c77371fee450b80e49c4a58d91342)	tags1 = spark.read.format("csv").option("header... load at NativeMethodAccessorImpl.java:0	2022/05/22 11:33:13	6 s	1/1	8/8
9 (6837360152381305406_6060119775019601837_1e2c77371fee450b80e49c4a58d91342)	tags1 = spark.read.format("csv").option("header... load at NativeMethodAccessorImpl.java:0	2022/05/22 11:33:13	0.5 s	1/1	1/1
8 (6837360152381305406_8606716216765786662_1e2c77371fee450b80e49c4a58d91342)	tags1 = tags.toPandas() y_hc = hc.fit_predict(t... wrapper at <command-244203776416437>:1	2022/05/22 11:31:01	1 s	1/1	8/8

Fig. 9 Spark jobs executing parallelly in multiple pools

```python
def recommendation_engine(userId, movieClustersFavouriteGenre, userClusterFavouriteGenre, userIdClusters):
    users_cluster = userIdClusters[userId]
    flag = 0
    for i in range(1, len(movieClustersFavouriteGenre) + 1):
        if movieClustersFavouriteGenre[i] == userClusterFavouriteGenre[users_cluster]:
            return i
    if flag == 0:
        r = random.randint(1, 35)
        return r
```

Step 4: Clustering user's with similar ratings.

```
from sklearn.metrics.pairwise import cosine_similarity

user_similarity = cosine_similarity(user_data)
user_similarity[np.isnan(user_similarity)] = 0
print(user_similarity)
print(user_similarity.shape)
```

Evaluation Of USER-USER

Cmd 30

```
test_features = X_test
test_features
```

```
user_final_ratings = np.multiply(user_predicted_ratings, dummy_train)
user_final_ratings.head()
```

Step 5: Content-based filtering on movie and user clusters.

The approach we followed is to:

1. Identify the movies present in the corresponding cluster obtained from CURE [5].
2. Identify the genres of the movies present in the given cluster.
3. Store the key value pairs of the cluster ID's and movie ID's and identify the genres present in the cluster.
4. Find the dominant genre present in the cluster.

Repeat Step 1 to 4 for rest of the clusters.

The user clusters obtained consist of ID's of the movies belonging to the cluster. We selected the movies which are favorite to the users, i.e., ratings >= 3.

The approach we followed is to:

1. Identify the users present in the corresponding cluster obtained from CURE.
2. Identify the genres of the movies favorite to the users present in the cluster.
3. Store the key value pairs of the cluster ID's and movie ID's and identify the genres present in the cluster.
4. Find the dominant genre present in the cluster.
5. Repeat step 1 to 4 for rest of the clusters.

Step 6: Calculating the RMSE using the predicted ratings of the user to the test movies dataset by splitting into train and split.

```
# Root Mean Squared Error
diff_sqr_matrix = (test - pred)**2
sum_of_squares_err = diff_sqr_matrix.sum().sum() # df.sum().sum() by default ignores null values

rmse = np.sqrt(sum_of_squares_err/total_non_nan)
print(rmse)
```

```
0.989180457090977
```

4 Conclusion and Future Work

Databricks allow to visualize the jobs running in parallel using the spark UI. Multiple jobs are running parallelly during the clustering process. After the previous process, we obtain a predicted cluster from the recommendation engine. All the movies present in the predicted cluster will be the movies which are to be recommended to the user (Fig. 9).

In present generation, there is enormous amount of data present, and we have variety of recommendation algorithms running specific to the area of interest and domain of the problem.

We can increase the effectiveness of the recommendations by using new variants, i.e., hybrid recommendation algorithms and come up with new and efficient algorithm.

The algorithm which we came up with is one such hybrid algorithm to provide proper and reliable personalized recommendations by considering factual data.

However, it is inevitable that the data is large and needs to be handled in a faster way. So, the idea of parallelization helped us in reducing the data into various smaller chunks and process them parallelly and utilize the resources effectively. Since we used spark distributed environment, this project can be scaled up to real world by using the AWS S3 service and process data in real time.

References

1. Das AS, Datar M, Garg A, Rajaram S (2007) Google news personalization: scalable online collaborative filtering. WWW
2. Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial database with noise. In: International conference on knowledge discovery in databases and data mining (KDD-96), Portland, Oregon, Aug 1996
3. Ester M, Kriegel H-P, Xu X (1995) A database interface for clustering in large spatial databases. In: International conference on knowledge discovery in databases and data mining (KDD-95), Montreal, Canada, Aug 1995
4. Guha S, Rastogi R, Shim K (1997) CURE: a clustering algorithm for large databases. Technical report, Bell Laboratories, Murray Hill
5. Drineas P, Frieze AM, Kannan R, Vempala S, Vinay V (2004) Clustering large graphs via the singular value decomposition (PDF). Machine Learn 56(1–3):9–33. doi:https://doi.org/10.1023/b:mach.0000033113.59016.96 (S2CID 5892850)
6. https://grouplens.org/datasets/movielens/20m/
7. Beckmann N, Kriegei H-P, Schneider R, Seeger B (1990) The R*-tree: an efficient and robust access method for points and rectangles. In: Proceedings of ACM SIGMOD, pp 322–331, Atlantic City, NJ, May 1990
8. Cormen TH, Leiserson CE, Rivest RL (1990) Introduction to algorithms. The MIT Press, Massachusetts